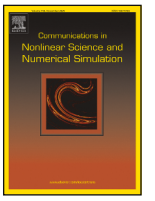
Commun Nonlinear Sci Numer Simulat 162 (2026) 110407

Contents lists available at ScienceDirect

Communications in Nonlinear Science and Numerical Simulation

journal homepage: www.elsevier.com/locate/cnsns



Research paper

Tube-type flexible performance preset boundary-guided reinforcement learning control for input saturated nonlinear systems

Ziyi Liu [a], Yangang Yao [a,*], Yu Kang [b,*], Yunbo Zhao [b], Jieqing Tan [c], Lichuan Gu [a,*], Qiang Li [a]

[a] School of Artificial Intelligence, Anhui Agricultural University, Hefei, 230036, China
[b] Department of Automation, University of Science and Technology of China, Hefei, 230027, China
[c] School of Mathematics, Hefei University of Technology, Hefei, 230601, China

ARTICLE INFO

Keywords:
Tube-type flexible performance preset boundary
Input saturated nonlinear systems
Reinforcement learning

ABSTRACT

This study presents a tube-type flexible performance preset boundary-guided reinforcement learning (TyFPPB-RL) control algorithm for input saturated nonlinear systems (ISNSs). By constructing a pair of TyFPPBs, a tube-type flexible prescribed performance control (TyFPPC) approach is proposed. Unlike traditional PPC (TPPC) algorithms, the significant advantage of the presented TyFPPC algorithm is its capacity to integrate multiple performance metrics, such as overshoot, steady-state accuracy, and convergence time. Meanwhile, the TyFPPC algorithm achieves the balance between performance requirements and input security through the design of a supplementary system. Furthermore, the RL strategy with the identifier-critic-actor structure (ICAS) is utilized to reduce the control system's cost function while achieving control performance, thus realizing system optimization. The effectiveness of the proposed algorithm is confirmed via simulation experiments.

1. Introduction

In the past few years, the control of constrained systems has received increasing attention and extensive research. Systems constraints can be classified into two distinct categories: one is performance constraints, such as system control accuracy and convergence time, and the other is physical constraints, such as system state constraints and input constraints. These constraints work in conjunction to ensure the system's long-term stable operation and to achieve the expected performance. For physical constraints, researchers have developed a variety of effective control algorithms [1–5]. In order to further enhance system control performance, Bechlioulis et al. [6] pioneered the PPC algorithm, which achieves the desired control objectives by persistently confining the tracking error within PPBs. Subsequently, the PPC algorithm was extended to various systems [7–9]. However, the limitation of the above TPPC algorithms, lies in that they can only ensure the achieving control objectives within an infinite time horizon, which restricts their practical applicability.

Practical engineering systems generally necessitate the attainment of specified control objectives within a finite time frame. For this purpose, the finite-time control (FTC) has become an effective alternative due to its properties such as rapid convergence, high

* Corresponding authors.
E-mail addresses: 24721982@stu.ahau.edu.cn (Z. Liu), ygyao@ustc.edu.cn (Y. Yao), kangduyu@ustc.edu.cn (Y. Kang), ybzhao@ustc.edu.cn (Y. Zhao), tjq20202020@126.com (J. Tan), glc@ahau.edu.cn (L. Gu), seuliqiang@ahau.edu.cn (Q. Li).

https://doi.org/10.1016/j.cnsns.2026.110407
Received 2 December 2025; Received in revised form 31 March 2026; Accepted 15 June 2026
Available online 23 June 2026

control accuracy, and robust anti-interference capability [10–18]. However, the settling time of FTC algorithms depends on initial conditions, or multiple control parameters, and the control accuracy is difficult to predict in advance. To achieve precise control, researchers innovatively integrated PPC with FTC [19,20], but this has inevitably increased controller design complexity. As a result, Zhao et al. [21] proposed a finite-time PPC (FTPPC) method, which ensures the tracking error attains the preset accuracy within finite time by designing a piecewise continuous PPBs. Then, the FTPPC was extended and applied to various systems [22–25]. However, the above TPPC and FTPPC algorithms exhibit a limitation, i.e., the PPBs of which are defined by a set of symmetric, strictly monotonic, or piecewise monotonic functions, which could result in unfavorable behaviors like chattering and overshoot during the initial phase. To tackle these challenges, some novel enhancing PPC (EPPC) algorithms that comprehensively account for both steady and transient-state performance were proposed [26–30]. It should be noted that the development of EPPC approaches relies on the assumption that the control input can be infinitely increased, without taking into account the input saturation that is commonly present in actual engineering, which is obviously contrary to the requirements of actual engineering systems.

In addition to the preset performance, input security is one of the important conditions for ensuring the long-term stable operation of the system. Accordingly, based on the transformation idea, the PPC for ISNSs were proposed [31,32]. Notably, these methods have a common limitation that they address performance requirements and input security in isolation. However, there is a significant mutual influence between the two: overly strict performance constraints can lead to input saturation, thereby triggering safety hazards; conversely, overly loose input constraints make it difficult to meet the performance indicators. To mitigate this conflict, Yong et al. [33] introduced a supplementary system to effectively balance performance preset with input security. Subsequently, this method was further extended to various system [34–37]. However, the above algorithms mostly focus on ensuring system performance and operational security, but do not consider the optimization of overall performance.

It is worth noting that energy conservation is particularly crucial for the long-term operation of industrial systems. Against this backdrop, optimal control achieves a balance between performance and resources by determining optimal control laws for dynamic systems, enabling them to minimize cost functions while satisfying performance requirements. To name just a few, Xu et al. [38] achieved optimal control by solving the optimal solution to the Hamilton-Jacobi-Bellman (HJB) equation. However, due to the strong nonlinear property, it is challenging to derive HJB's analytical solution. To overcome this challenge, leveraging the adaptive approximation capability of RL to achieve optimal control is an effective approach, as it avoids the need to directly solve the HJB equation. For instance, the real-time collaborative training of control actions and performance indicators is achieved via the RL algorithm with an actor-critic structure [39,40]. Subsequently, Wen et al. [41] addressed the unknown dynamic characteristics of multi-agent systems by designing a novel backstepping control framework integrated with RL. Bai et al. [42,43] presented the RL-based algorithms for nonstrict-feedback discrete-time systems. Zhu et al. [44] presented an ICAS-based RL control approach for stochastic systems, where an adaptive identifier is employed to approximate uncertain nonlinear terms. Furthermore, the methods that integrated the PPC with RL were presented for stochastic systems [45], ISNSs [46], MAS [47], as well as USV [48]. Despite notable advancements in RL-based PPC, the RL control with high performance (i.e., not only consider the preset of multiple performance indicators such as overshoot, steady-state accuracy and convergence time, and the balance between them and input safety, but also taken into account the optimization of global performance) has received little attention.

Building on the preceding analysis, this study introduces a TyFPPB-guided RL control for ISNSs. The main research results are as follows:

1. Unlike the existing TPPC [6–9] and FTPPC [21–25] algorithms, the PPBs of which are composed of a set of symmetric, strictly monotonic, or piecewise monotonic functions, which could result in unfavorable behaviors like chattering and overshoot during the initial phase. By designing a pair of TyFPPBs, a novel TyFPPC approach is proposed, its significant advantage is that the presented TyFPPC algorithm can comprehensively consider multiple performance indicators, including the overshoot, convergence time and steady state accuracy. Furthermore, by designing a supplementary system, the presented TyFPPC method also achieves the balance between performance requirements and input security.
2. By integrating the TyFPPC with the RL strategy, this paper first proposes a TyFPPB-RL optimal control method for ISNSs based on the ICAS. The designed TyFPPB-RL optimal control approach can minimize the cost function of the control system while satisfying performance constraints and input constraints, thereby achieving the overall optimization of the system.

2. Problem formulation and preliminaries

2.1. Problem formulation

Consider the ISNSs below

$$\begin{cases} \dot{\chi}_j(t) = \chi_{j+1} + f_j(\bar{\chi}_j) \\ \dot{\chi}_n(t) = S(u) + f_n(\bar{\chi}_n) \\ y = \chi_1 \end{cases} \tag{1}$$

with $j = 1, 2, \ldots, n-1$. $\bar{\chi}_j = [\chi_1, \chi_2, \ldots, \chi_j]^{\mathrm{T}}$ and $\bar{\chi}_n = [\chi_1, \chi_2, \ldots, \chi_n]^{\mathrm{T}}$ denote the state vector, the system output is denoted by y, $f_j(\cdot)$ represents a nonlinear continuous function with uncertainty. The actual control input $S(u)$ satisfy

$$S(u) = \begin{cases} \mathrm{sgn}(u)u_s, & |u| > u_s \\ u, & |u| \le u_s \end{cases} \tag{2}$$

where u_s denotes the maximum permissible threshold. Adopting the estimation below to handle the acute angles of $S(u)$

$$S(u) = s_1(u) + s_2(u) \tag{3}$$

here $s_2(u) = S(u) - s_1(u)$ with $|s_2(u)| \le u_s(1 - \tanh(1)) = \bar{s}_2$, $s_1(u) = u_s \tanh\left(u/u_s\right)$.

This paper proposes an innovative TyFPPB-RL control strategy designed to ensure semi-global uniform ultimate boundedness (SGUUB) of all closed-loop system signals, while rigorously confining the tracking error within the TyFPPB.

The algorithm introduced in this study rests upon a logically sound premise

Assumption 1 ([3,5]). The trajectory used as a reference y_r and its respective derivatives $y_r^{(j)}$ are bounded and continuous.

2.2. Radial basis function neural network (RBFNN)

Lemma 1 ([20]). *The unknown nonlinear term $\xi(\boldsymbol{X})$ is capable of estimated via a RBFNN in the form shown below*

$$\xi(\boldsymbol{X}) = \epsilon(\boldsymbol{X}) + \boldsymbol{w}^{\mathrm{T}}\boldsymbol{S}(\boldsymbol{X}), \quad (|\epsilon(\boldsymbol{X})| \le \epsilon, \epsilon \in \boldsymbol{R}^{+}) \tag{4}$$

where $\epsilon(\boldsymbol{X})$, $\boldsymbol{w}$, $\boldsymbol{S}(\boldsymbol{X})$ and $\boldsymbol{X}$ denote the approximation error, weight vector, basis function vector and input, $S(\boldsymbol{X}) = [s_1(\boldsymbol{X}), \ldots, s_m(\boldsymbol{X})]^{\mathrm{T}}$ with $m \in \boldsymbol{R}^{+}$, and

$$s_i(\boldsymbol{X}) = \exp\left(\frac{-\left(\boldsymbol{X} - \boldsymbol{k}_i\right)^{\mathrm{T}}\left(\boldsymbol{X} - \boldsymbol{k}_i\right)}{\varkappa_i^{\,2}}\right) \tag{5}$$

here, $\varkappa_i$ represents the width of the ith neural node, while $\boldsymbol{k}_i$ stands for its corresponding center vector.

3. Main results

3.1. Tube-type flexible performance preset boundary

This paper requires the tracking error $e_1(t) = \chi_1 - y_r$ (with y_r representing the reference signal) to satisfy the following constraint

$$\mathcal{P}_l < e_1(t) < \mathcal{P}_h \tag{6}$$

where $\left[\mathcal{P}_l, \mathcal{P}_h\right]^{\mathrm{T}}$ denotes the TyFPPB as follows

$$\left[\mathcal{P}_l, \mathcal{P}_h\right]^{\mathrm{T}} = \left[p_l, p_h\right]^{\mathrm{T}} + \tanh(\varrho(t))\boldsymbol{\Gamma}_1 \tag{7}$$

where $\left[p_l, p_h\right]^{\mathrm{T}} = \rho(t)\boldsymbol{\Gamma}_2 + \mathrm{sgn}(e_1(0))\left(\rho(t) - \rho_T\right)\mathbf{I}$, $\mathbf{I} = [1,1]^{\mathrm{T}}$ and $\boldsymbol{\Gamma}_i = [-\gamma_i, \gamma_i]^{\mathrm{T}}$ with $i = 1, 2$, $\gamma_1 \in \boldsymbol{R}^{+}, \gamma_2 \in [0,1]$. Additionally

$$\rho(t) = \begin{cases} \mathrm{csch}\left(\rho_0 + \frac{lt}{T-t}\right) + \rho_T, & 0 \le t < T \\ \rho_T, & t \ge T \end{cases} \tag{8}$$

with $\rho_0, \rho_T, T, l \in \boldsymbol{R}^{+}$ be the design parameters.

$\varrho(t)$ represents the output generated by the following supplementary system [26]

$$\dot{\varrho}(t) = -b_1\varrho(t) + b_2(\tau_1(t) + \tau_2(t)), \varrho(0) = 0 \tag{9}$$

where $b_1, b_2 \in \boldsymbol{R}^{+}$, $\tau_1(t) = \left(u - u_s\right)\left[\mathrm{sgn}(u - u_s) + 1\right]$, $\tau_2(t) = \left(u + u_s\right)\left[\mathrm{sgn}(u + u_s) - 1\right]$.

Remark 1. It is not difficult to conclude from the definition of $\tau_i(t)$ $(i = 1, 2)$ that $\sum_{i=1}^{2}\tau_i(t) > 0$, and if and only if $|u| > u_s$. Furthermore, in conjunction with the designed supplementary system (9), it further follows that if and only if $|u| > u_s$, $\varrho(t) > 0$ holds. According to (7), It is evident that if $|u| > u_s$, i.e., the control signal surpasses its upper allowable limit, the TyFPPB undergoes adaptive expansion, and once $|u| \le u_s$, $\left[\mathcal{P}_l, \mathcal{P}_h\right]^{\mathrm{T}} = \left[p_l, p_h\right]^{\mathrm{T}}$ holds, meaning that the TyFPPB adaptively reverts to the initial PPB. In other words, the design of supplementary system (9) achieves the balance between performance requirements and input security.

Remark 2. The existing TPPC algorithms [6–9] and FTPPC algorithms [21–25] have a common limitation, i.e., the PPBs of which are composed of a set of symmetric strictly monotonic or piecewise monotonic functions, which may induce chattering, overshoot, and other undesirable phenomena in the initial stage. Although the EPPC methods comprehensively consider both transient and steady-state performance, they all ignore the input constraints. The advantage of the proposed TyFPPC lies in not only inheriting the advantages of the EPPC methods, but also further considering the intrinsic connection between input saturation and performance preset, and establishing the balance between input safety and performance requirements. To clearly illustrate this improvement, Fig. 1 presents the error curves under different PPBs.

- As can be seen from subfigure (a), the PPB of TPPC is strictly monotonous, which means that only asymptotic convergence can be guaranteed. Neither the steady-state precision nor the speed of convergence can meet the specified requirements.
- As can be seen from subfigure (b),the FTPPC can achieve the preset steady-state accuracy and convergence time, while the PPB of which presents a “trumpet” shape, so the transient performance before the preset time T is difficult to guarantee.

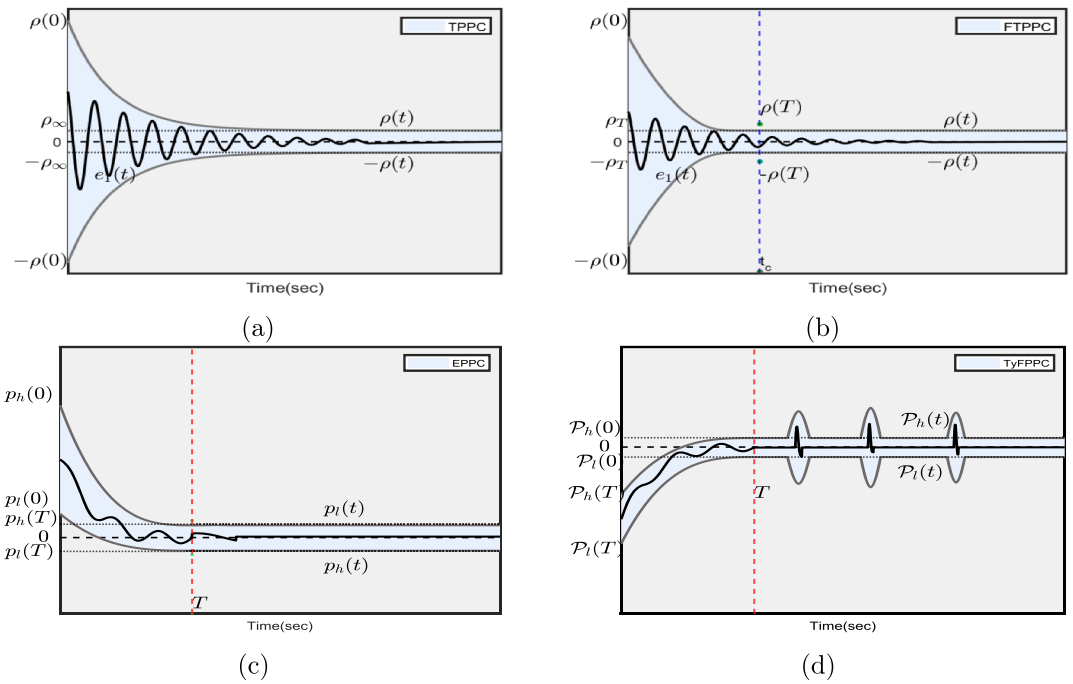


Fig. 1. The tracking error curves under four different PPBs.

- Subgraph (c) shows the PPB of the EPPC algorithm, unlike the TPPC and FTPPC algorithms, the EPPC comprehensively account for steady-state accuracy, convergence time, and overshoot, thereby achieving better control performance than those of TPPC and FTPPC, the input constraints, however, were ignored.
- Subgraphs (d) shows the constructed TyFPPB, apparently, the proposed TyFPPC not only inheriting the advantages of the EPPC methods, but also further considering the intrinsic connection between performance preset and input saturation, i.e., when the control input exceeds the maximum threshold, the TyFPPB undergoes adaptive expansion, and once $|u| \leq u_s$, the TyFPPB adaptively reverts to the initial PPB, meaning that the proposed TyFPPC algorithm also achieves the balance between performance requirements and input security.

3.2. TyFPPB-guided RL control design

The error transformation is defined as follows

$$\varsigma(t) = \ln\left(\frac{\delta(t)}{1-\delta(t)}\right) \tag{10}$$

with $\delta(t) = \big(e_1(t) - \mathcal{P}_l(t)\big)/\big(\mathcal{P}_h(t) - \mathcal{P}_l(t)\big)$.

Using the definition of $\delta(t)$, it follows that if $e_1(0) \in \big(\mathcal{P}_l(0), \mathcal{P}_h(0)\big)$, then $\delta(0) \in (0,1)$. From (10), it is further inferred that $\varsigma(t) \to \infty$ if and only if $\delta(t) \to 1^-$ or $\delta(t) \to 0^+$. This indicates that, as long as $0 < \delta(0) < 1$ and $\varsigma(t)$ is bounded, $0 < \delta(t) < 1$ holds, which further follows that

$$\mathcal{P}_l(t) < e_1(t) < \mathcal{P}_h(t), \forall t \geq 0. \tag{11}$$

In the subsequent sections, to simplify matters, let $\mathcal{P}_l(t) = \mathcal{P}_l$, $\mathcal{P}_h(t) = \mathcal{P}_h$, $\delta(t) = \delta$ and $\varsigma(t) = \varsigma$.

Let

$$\begin{cases} z_1 = \varsigma, \\ z_j = \chi_j - \hat{\alpha}^*_{j-1}, j = 2:n-1, \\ z_n = \chi_n - \hat{\alpha}^*_{n-1} - o \end{cases} \tag{12}$$

where α_j and o will be given later.

Step 1: From (1), (10) and (12), one obtain

$$\dot{z}_1 = \varphi \dot{e}_1 + \beta \tag{13}$$

where $\varphi = 1/\delta(1-\delta)(\mathcal{P}_h - \mathcal{P}_l)$,$\beta = \varphi[\dot{\mathcal{P}}_h(\mathcal{P}_l - e_1) + \dot{\mathcal{P}}_l(e_1 - \mathcal{P}_h)]/\big(\mathcal{P}_h - \mathcal{P}_l\big)$.

Let α_1 serve as the virtual controller, α_1^* stands for its optimal counterpart. Design the optimal performance index function (OPIF) in the following way

$$\begin{aligned}\mathcal{J}_1^*(z_1) &= \min_{\alpha_1\in\Psi(\Omega)}\left(\int_t^\infty C_1\big(\alpha_1(z_1), z_1(s)\big)\mathrm{d}s\right)\\ &= \int_t^\infty C_1\big(\alpha_1^*(z_1), z_1(s)\big)\mathrm{d}s\end{aligned} \tag{14}$$

where $C_1(\alpha_1, z_1) = z_1^2 + \alpha_1^2$ serves as the cost function,Ω stands for a compact set with the origin included.

Treat χ_2 as α_1^*. Based on (13)-(14), the HJB equation is obtained

$$\mathbb{H}_1\left(z_1, \alpha_1^*, \frac{\mathrm{d}\mathcal{J}_1^*}{\mathrm{d}z_1}\right) = z_1^2 + \alpha_1^{*2} + \frac{\mathrm{d}\mathcal{J}_1^*}{\mathrm{d}z_1}\dot{z}_1 = 0. \tag{15}$$

By solving $\partial\mathbb{H}_1/\partial\alpha_1^* = 0$, one can derive

$$\alpha_1^* = -\frac{\varphi}{2}\frac{\mathrm{d}\mathcal{J}_1^*(z_1)}{\mathrm{d}z_1}. \tag{16}$$

Step j ($j = 2 : n-1$) : Combine (1) and (12), it is possible to derive

$$\dot{z}_j = \chi_{j+1} + f_j(\bar{\chi}_j) - \dot{\hat{\alpha}}_{j-1}^*. \tag{17}$$

Let α_j serve as the optimal virtual control law, and α_j^* stands for its optimal counterpart, thus define the OPIF as follows

$$\begin{aligned}\mathcal{J}_j^*(z_j) &= \min_{\alpha_j\in\Psi(\Omega)}\left(\int_t^\infty C_j\big(\alpha_j(z_j), z_j(s)\big)\mathrm{d}s\right)\\ &= \int_t^\infty C_j\Big(\alpha_j^*(z_j), z_j(s)\Big)\mathrm{d}s\end{aligned} \tag{18}$$

where the cost function is defined as $C_j(\alpha_j, z_j) = \alpha_j^2 + z_j^2$.

When χ_{j+1} is taken as α_j^*, we can derive the HJB equation in the following manner

$$\mathbb{H}_j\left(z_j, \alpha_j^*, \frac{\mathrm{d}\mathcal{J}_j^*}{\mathrm{d}z_j}\right) = z_j^2 + \alpha_j^{*2} + \frac{\mathrm{d}\mathcal{J}_j^*}{\mathrm{d}z_j}\dot{z}_j = 0. \tag{19}$$

Solving $\partial\mathbb{H}_j/\partial\alpha_j^* = 0$, yields the optimal control α_j^* expressed as

$$\alpha_j^* = -\frac{1}{2}\frac{\mathrm{d}\mathcal{J}_j^*(z_j)}{\mathrm{d}z_j}. \tag{20}$$

Step n: Construct a supplementary system to counteract the adverse effects caused by input constraint

$$\dot{o} = \big(s_1(u) - u\big) - o. \tag{21}$$

By integrating (1),(3),(12) and (21), the desired result follows

$$\dot{z}_n = f_n(\bar{\chi}_n) + s_2(u) + u - \dot{\hat{\alpha}}_{n-1}^* + o. \tag{22}$$

Taking u^* as the optimal controller, we proceed to define the OPIF as

$$\begin{aligned}\mathcal{J}_n^*(z_n) &= \min_{u\in\Psi(\Omega)}\left(\int_t^\infty C_n\big(u(z_n), z_n(s)\big)\mathrm{d}s\right)\\ &= \int_t^\infty C_n\big(u^*(z_n), z_n(s)\big)\mathrm{d}s\end{aligned} \tag{23}$$

where $C_n(u, z_n) = z_n^2 + u^2$ is a cost function.

The HJB equation corresponding to (23) can be written as

$$\mathbb{H}_n\left(z_n, u^*, \frac{\mathrm{d}\mathcal{J}_n^*}{\mathrm{d}z_n}\right) = z_n^2 + u^{*2} + \frac{\mathrm{d}\mathcal{J}_n^*}{\mathrm{d}z_n}\dot{z}_n = 0. \tag{24}$$

Solving $\partial\mathbb{H}_n/\partial u^* = 0$ gives the optimal control u^* as follows

$$u^* = -\frac{1}{2}\frac{\mathrm{d}\mathcal{J}_n^*(z_n)}{\mathrm{d}z_n}. \tag{25}$$

Owing to the uncertainties in $\mathrm{d}\mathcal{J}_j^*(z_j)/\mathrm{d}z_j$, this term becomes infeasible for practical use. To overcome this limitation, this section proposes an optimal controller based on the ICAS, i.e., the identifier, critic and actor modules are responsible for estimating unknown nonlinear dynamics, evaluating the optimal cost function, and implementing control policies, respectively.

Decompose $\mathrm{d}\mathcal{J}_j^*(z_j)/\mathrm{d}z_j$ into the following form

$$\frac{\mathrm{d}\mathcal{J}_1^*}{\mathrm{d}z_1} = \frac{2v_1z_1}{\varphi^2} + \frac{2\xi_1\left(\boldsymbol{X}_{\xi_1}\right)}{\varphi^2} + \frac{\mathcal{J}_1^0\left(\boldsymbol{X}_{\mathcal{J}_1}\right)}{\varphi^2}, \tag{26}$$

$$\frac{\mathrm{d}\mathcal{J}_j^*}{\mathrm{d}z_j} = 2v_jz_j + 2\xi_j\left(\boldsymbol{X}_{\xi_j}\right) + \mathcal{J}_j^0\left(\boldsymbol{X}_{\mathcal{J}_j}\right), j = 2, \ldots, n \tag{27}$$

where $v_j \in \boldsymbol{R}^+$, $\xi_1\left(\boldsymbol{X}_{\xi_1}\right) = \varphi\left(f_1(\chi_1) - \dot{y}_r + \beta/\varphi + \varphi z_1/2 + 3z_1/4\right)$, $\xi_j\left(\boldsymbol{X}_{\xi_j}\right) = f_j\left(\bar{\chi}_j\right) - \dot{\hat{\alpha}}_{j-1}^* + 7z_j/4, j = 2, \ldots, n-1$, $\xi_n\left(\boldsymbol{X}_{\xi_n}\right) = f_n\left(\bar{\chi}_n\right) - \dot{\hat{\alpha}}_{n-1}^* + 7z_j/4 + o$,$\mathcal{J}_1^o(\boldsymbol{X}_{\mathcal{J}_1}) = -2v_1z_1 - 2\xi_1(\boldsymbol{X}_{\xi_1}) + \varphi^2\mathrm{d}\mathcal{J}_1^*/\mathrm{d}z_1$, $\mathcal{J}_j^o(\boldsymbol{X}_{\mathcal{J}_j}) = -2v_jz_j - 2\xi_j(\boldsymbol{X}_{\xi_j}) + \mathrm{d}\mathcal{J}_j^*/\mathrm{d}z_j, j = 2, \ldots, n$,$\boldsymbol{X}_{\xi_1} = \left[\bar{\chi}_j, y_r\right]^{\mathrm{T}}$, $\boldsymbol{X}_{\mathcal{J}_j} = \left[\bar{\chi}_j, y_r, z_j\right]^{\mathrm{T}}$.

Let $\alpha_n^* = u^*$, substituting (26)–(27) into (16), (20) and (25) yields

$$\alpha_1^* = -\frac{v_1z_1}{\varphi} - \frac{\xi_1\left(\boldsymbol{X}_{\xi_1}\right)}{\varphi} - \frac{\mathcal{J}_1^0\left(\boldsymbol{X}_{\mathcal{J}_1}\right)}{2\varphi}, \tag{28}$$

$$\alpha_j^* = -v_jz_j - \xi_j\left(\boldsymbol{X}_{\xi_j}\right) - \frac{\mathcal{J}_j^0\left(\boldsymbol{X}_{\mathcal{J}_j}\right)}{2}. \tag{29}$$

The RBFNNs are utilized to approximate $\xi_j\left(\boldsymbol{X}_{\xi_j}\right)$ and $\mathcal{J}_j^0\left(\boldsymbol{X}_{\mathcal{J}_j}\right)$ as detailed below

$$\xi_j\left(\boldsymbol{X}_{\xi_j}\right) = \boldsymbol{w}_{\xi_j}^{*\mathrm{T}}\boldsymbol{S}_{\xi_j}\left(\boldsymbol{X}_{\xi_j}\right) + \epsilon_{\xi_j}\left(\boldsymbol{X}_{\xi_j}\right), \tag{30}$$

$$\mathcal{J}_j^0\left(\boldsymbol{X}_{\mathcal{J}_j}\right) = \boldsymbol{w}_{\mathcal{J}_j}^{*\mathrm{T}}\boldsymbol{S}_{\mathcal{J}_j}\left(\boldsymbol{X}_{\mathcal{J}_j}\right) + \epsilon_{\mathcal{J}_j}\left(\boldsymbol{X}_{\mathcal{J}_j}\right) \tag{31}$$

here, $\boldsymbol{w}_{\xi_j}^*$ and $\boldsymbol{w}_{\mathcal{J}_j}^*$ denote the ideal weights, $\boldsymbol{S}_{\xi_j}(\boldsymbol{X}_j)$ and $\boldsymbol{S}_{\mathcal{J}_j}(\boldsymbol{X}_j)$ stand for the basis function vectors. The approximation errors are given by $\epsilon_{\xi_j}(\boldsymbol{X}_j)$ and $\epsilon_{\mathcal{J}_j}(\boldsymbol{X}_j)$, which satisfy $\epsilon_{\xi_j}(\boldsymbol{X}_j) < \bar{\epsilon}_{\xi_j}$, $\epsilon_{\mathcal{J}_j}(\boldsymbol{X}_j) < \bar{\epsilon}_{\mathcal{J}_j}$ with $\bar{\epsilon}_{\xi_j}, \bar{\epsilon}_{\mathcal{J}_j} \in \boldsymbol{R}^+$. For the sake of simplicity, in subsequent sections $\boldsymbol{S}_{\xi_j} = \boldsymbol{S}_{\xi_j}(\boldsymbol{X}_j), \boldsymbol{S}_{\mathcal{J}_j} = \boldsymbol{S}_{\mathcal{J}_j}(\boldsymbol{X}_j), \epsilon_{\xi_j} = \epsilon_{\xi_j}(\boldsymbol{X}_j), \epsilon_{\mathcal{J}_j} = \epsilon_{\mathcal{J}_j}(\boldsymbol{X}_j)$.

Substituting (30)–(31) into (26)–(29) allows us to derive

$$\frac{\mathrm{d}\mathcal{J}_1^*}{\mathrm{d}z_1} = \frac{2v_1z_1}{\varphi^2} + \frac{2\boldsymbol{w}_{\xi_1}^{*\mathrm{T}}\boldsymbol{S}_{\xi_1}}{\varphi^2} + \frac{\boldsymbol{w}_{\mathcal{J}_1}^{*\mathrm{T}}\boldsymbol{S}_{\mathcal{J}_1}}{\varphi^2} + \epsilon_1, \tag{32}$$

$$\frac{\mathrm{d}\mathcal{J}_j^*}{\mathrm{d}z_j} = 2v_jz_j + 2\boldsymbol{w}_{\xi_j}^{*\mathrm{T}}\boldsymbol{S}_{\xi_j} + \boldsymbol{w}_{\mathcal{J}_j}^{*\mathrm{T}}\boldsymbol{S}_{\mathcal{J}_j} + \epsilon_j, \tag{33}$$

$$\alpha_1^* = -\frac{v_1z_1}{\varphi} - \frac{\boldsymbol{w}_{\xi_1}^{*\mathrm{T}}\boldsymbol{S}_{\xi_1}}{\varphi} - \frac{\boldsymbol{w}_{\mathcal{J}_1}^{*\mathrm{T}}\boldsymbol{S}_{\mathcal{J}_1}}{2\varphi} - \frac{\varphi\epsilon_1}{2}, \tag{34}$$

$$\alpha_j^* = -v_jz_j - \boldsymbol{w}_{\xi_j}^{*\mathrm{T}}\boldsymbol{S}_{\xi_j} - \frac{\boldsymbol{w}_{\mathcal{J}_j}^{*\mathrm{T}}\boldsymbol{S}_{\mathcal{J}_j}}{2} - \frac{\epsilon_j}{2} \tag{35}$$

where $\epsilon_1 = 2\epsilon_{\xi_1}/\varphi^2 + \epsilon_{\mathcal{J}_1}/\varphi^2$,$\epsilon_j = 2\epsilon_{\xi_j} + \epsilon_{\mathcal{J}_j}$, $j = 2:n$.

Since the ideal weights $\boldsymbol{w}_{\xi_j}^*$ and $\boldsymbol{w}_{\mathcal{J}_j}^*$ are not available, the virtual control optimization, guided by RL, employs three NNs: an identifier network, a critic network, and an actor network.

A NN identifier is utilized for approximating $\xi_j\left(\boldsymbol{X}_{\xi_j}\right)$

$$\hat{\xi}_j\left(\boldsymbol{X}_{\xi_j}\right) = \hat{\boldsymbol{w}}_{\xi_j}^{\mathrm{T}}\boldsymbol{S}_{\xi_j} \tag{36}$$

where $\hat{\xi}_j\left(\boldsymbol{X}_{\xi_j}\right)$ indicates the output of the identifier, $\hat{\boldsymbol{w}}_{\xi_j}^T$ stands for the weight of the identifier NN, and $\boldsymbol{S}_{\xi_j}$ indicates the basis function vector.

$\hat{\boldsymbol{w}}_{\xi_j}$ tuned via the following adaptive law

$$\dot{\hat{\boldsymbol{w}}}_{\xi_j} = \boldsymbol{\Upsilon}_j\left(z_j\boldsymbol{S}_{\xi_j} - \eta_j\hat{\boldsymbol{w}}_{\xi_j}\right) \tag{37}$$

where $\boldsymbol{\Upsilon}_j$ represents a matrix that is positive definite, and $\eta_j \in \boldsymbol{R}^+$.

The effectiveness of the control strategy is assessed based on the critic NN described subsequently

$$\frac{\mathrm{d}\hat{\mathcal{J}}_1^*}{\mathrm{d}z_1} = \frac{2v_1z_1}{\varphi^2} + \frac{2\hat{\boldsymbol{w}}_{\xi_1}^{\mathrm{T}}\boldsymbol{S}_{\xi_1}}{\varphi^2} + \frac{\hat{\boldsymbol{w}}_{c_1}^{\mathrm{T}}\boldsymbol{S}_{\mathcal{J}_1}}{\varphi^2}, \tag{38}$$

$$\frac{\mathrm{d}\hat{\mathcal{J}}_j^*}{\mathrm{d}z_j} = 2v_jz_j + 2\hat{\boldsymbol{w}}_{\xi_j}^{\mathrm{T}}\boldsymbol{S}_{\xi_j} + \hat{\boldsymbol{w}}_{c_j}^{\mathrm{T}}\boldsymbol{S}_{\mathcal{J}_j} \tag{39}$$

where $\mathrm{d}\hat{\mathcal{J}}_j^*/\mathrm{d}z_j$ denotes the estimation of $\mathrm{d}\mathcal{J}_j^*/\mathrm{d}z_j$ and $\hat{\boldsymbol{w}}_{c_j}$ represents the weight of the critic NN.

$\hat{\boldsymbol{w}}_{c_j}$ follows the following weight update rules

$$\dot{\hat{\boldsymbol{w}}}_{c_j} = -L_{c_j}\boldsymbol{S}_{\mathcal{J}_j}\boldsymbol{S}_{\mathcal{J}_j}^{\mathrm{T}}\hat{\boldsymbol{w}}_{c_j} \tag{40}$$

where $L_{c_j} > 0$ denotes the parameter designed for the critic.

The construction of the actor NN proceeds as follows

$$\hat{\alpha}_1^* = -\frac{\upsilon_1 z_1}{\varphi} - \frac{\hat{\boldsymbol{w}}_{\xi_1}^{\mathrm{T}}\boldsymbol{S}_{\xi_1}}{\varphi} - \frac{\hat{\boldsymbol{w}}_{a1}^{\mathrm{T}}\boldsymbol{S}_{\mathcal{J}_1}}{2\varphi}, \tag{41}$$

$$\hat{\alpha}_j^* = -\upsilon_j z_j - \hat{\boldsymbol{w}}_{\xi_j}^{\mathrm{T}}\boldsymbol{S}_{\xi_j} - \frac{\hat{\boldsymbol{w}}_{a_j}^{\mathrm{T}}\boldsymbol{S}_{\mathcal{J}_j}}{2} \tag{42}$$

where $\hat{\alpha}_j^*$ represents the estimation of α_j^* and $\hat{\boldsymbol{w}}_{a_j}$ denotes the weight of the actor NN.

$\hat{\boldsymbol{w}}_{a_j}$ follows the following weight update rules

$$\dot{\hat{\boldsymbol{w}}}_{a_j} = -\boldsymbol{S}_{\mathcal{J}_j}\boldsymbol{S}_{\mathcal{J}_j}^{\mathrm{T}}\left(L_{a_j}\left(\hat{\boldsymbol{w}}_{a_j} - \hat{\boldsymbol{w}}_{c_j}\right) + L_{c_j}\hat{\boldsymbol{w}}_{c_j}\right) \tag{43}$$

where $L_{a_j} > 0$ represents the actor designed parameter.

The Bellman residual error $\bar{\mathbb{H}}_j(t)$ is defined here as follows

$$\begin{aligned}\bar{\mathbb{H}}_j(t) &= \mathbb{H}_j\left(z_j, \hat{\alpha}_j^*, \frac{\mathrm{d}\hat{\mathcal{J}}_j^*}{\mathrm{d}z_j}\right) - \mathbb{H}_j\left(z_j, \alpha_j^*, \frac{\mathrm{d}\mathcal{J}_j^*}{\mathrm{d}z_j}\right)\\ &= \mathbb{H}_j\left(z_j, \hat{\alpha}_j^*, \frac{\mathrm{d}\hat{\mathcal{J}}_j^*}{\mathrm{d}z_j}\right).\end{aligned} \tag{44}$$

From the preceding analysis, the optimized solution $\hat{\alpha}_j^*$ should satisfy $\bar{\mathbb{H}}_j(t) = \mathbb{H}_j\left(z_j, \hat{\alpha}_j^*, \mathrm{d}\hat{\mathcal{J}}_j^*/\mathrm{d}z_j\right) \to 0$. Provided that $\mathbb{H}_j\left(z_j, \hat{\alpha}_j^*, \mathrm{d}\hat{\mathcal{J}}_j^*/\mathrm{d}z_j\right) = 0$ holds true and has a unique solution, the equivalence below is valid

$$\frac{\partial\mathbb{H}_1\left(z_1, \hat{\alpha}_1^*, \frac{\mathrm{d}\hat{\mathcal{J}}_1^*}{\mathrm{d}z_1}\right)}{\partial\hat{\boldsymbol{w}}_{a_1}} = \frac{1}{2\varphi^2}\boldsymbol{S}_{\mathcal{J}_1}^{\mathrm{T}}\boldsymbol{S}_{\mathcal{J}_1}\left(\hat{\boldsymbol{w}}_{a_1} - \hat{\boldsymbol{w}}_{c_1}\right) = 0, \tag{45}$$

$$\frac{\partial\mathbb{H}_j\left(z_j, \hat{\alpha}_j^*, \frac{\mathrm{d}\hat{\mathcal{J}}_j^*}{\mathrm{d}z_j}\right)}{\partial\hat{\boldsymbol{w}}_{a_j}} = \frac{1}{2}\boldsymbol{S}_{\mathcal{J}_j}^{\mathrm{T}}S_{\mathcal{J}_j}\left(\hat{\boldsymbol{w}}_{a_j} - \hat{\boldsymbol{w}}_{c_j}\right) = 0. \tag{46}$$

To derive RL update laws that ensure (45)-(46) are satisfied, the subsequent positive-definite function is established

$$M_j(t) = \left((\hat{\boldsymbol{w}}_{a_j} - \hat{\boldsymbol{w}}_{c_j})^{\mathrm{T}}(\hat{\boldsymbol{w}}_{a_j} - \hat{\boldsymbol{w}}_{c_j})\right). \tag{47}$$

Obviously, $M_j(t) = 0$ is equivalent to (45)-(46). The construction of the update laws in (40) and (43) follows from the subsequent design principles. Considering that $\partial M_j(t)/\partial\hat{\boldsymbol{w}}_{a_j}(t) = -\partial M_j(t)/\partial\hat{\boldsymbol{w}}_{c_j}(t) = 2\left(\hat{\boldsymbol{w}}_{a_j}(t) - \hat{\boldsymbol{w}}_{c_j}(t)\right)$, determining the derivative with respect to time of $M_j(t)$ along (40) and (43) results in

$$\frac{\mathrm{d}M_j(t)}{\mathrm{d}t} = \frac{\partial M_j(t)}{\partial\hat{\boldsymbol{w}}_{a_j}}\dot{\hat{\boldsymbol{w}}}_{a_j} + \frac{\partial M_j(t)}{\partial\hat{\boldsymbol{w}}_{c_j}}\dot{\hat{\boldsymbol{w}}}_{c_j} \le 0. \tag{48}$$

Inequality (48) ensures $M_j(t) = 0$ via update laws (40) and (43), thereby guaranteeing that conditions (45) and (46) are satisfied.

Remark 3. Note that in the nth step, what this paper optimizes is indeed the design control input u, rather than the actual control input $S(u)$. The design concept is to first minimize the designed control input. Of course, the minimized result cannot guarantee that its optimized value is less than the saturation threshold. Therefore, before inputting the plant, saturation processing was carried out. The resulting outcome is actually similar to optimizing the actual control input.

Remark 4. Unlike the existing PPC approaches [6–9,21–26], in addition to the significant advantages of the designed TyFPPB (see Remarks 1-2), another advantage of the proposed approach lies in the integration of the TyFPPB with the RL strategy, and a TyFPPB-RL optimal control method for ISNSs based on the ICAS is first presented in this article. The designed TyFPPB-RL optimal approach can minimize the cost function of the control system while satisfying performance constraints and input constraints, thereby achieving the overall optimization of the system. Fig. 2 presents the TyFPPB-guided RL control framework under ICAS.

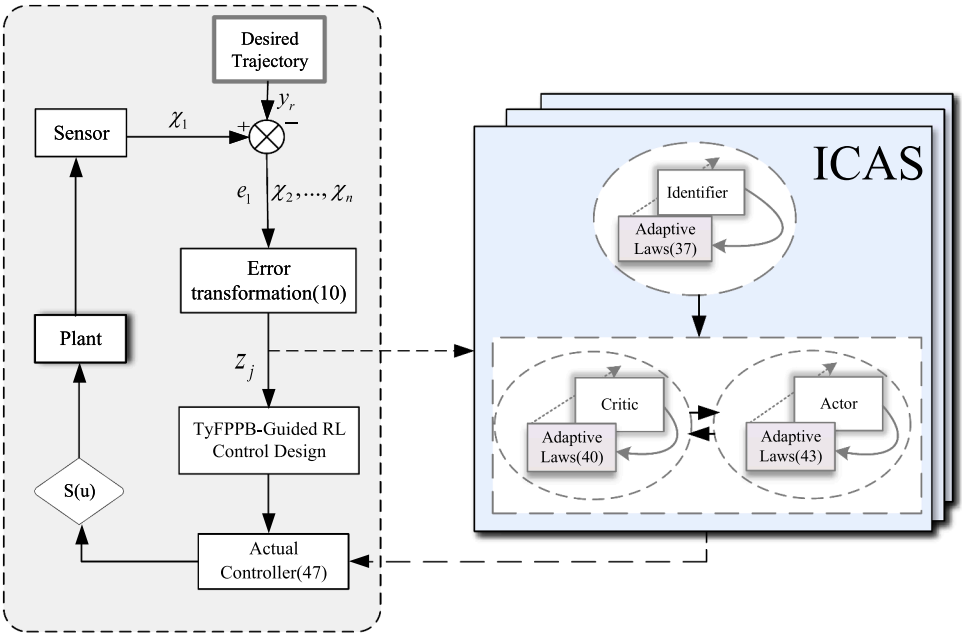


Fig. 2. The TyFPPB-guided RL control framework under ICAS.

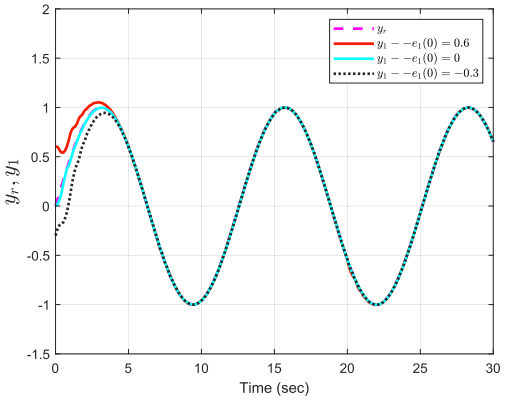


Fig. 3. Tracking trajectories under three distinct errors.

3.3. Stability analysis

Theorem 1. *Considering ISNSs (1) that satisfy Assumption 1, and the initial condition $\mathcal{P}_l \le e_1(0) \le \mathcal{P}_h$, the presented TyFPPB-guided RL control algorithm can guarantee that*

1) all signals of the closed-loop system (1) are SGUUB,
2) the system output successfully tracks the reference signals while maintaining the tracking error within the TyFPPB.

Proof. Choose the LF candidate

$$V = \sum_{j=1}^{n} \frac{z_j^2 + \tilde{\boldsymbol{w}}_{\xi_j}^{\mathrm{T}} \boldsymbol{\Upsilon}_j^{-1} \tilde{\boldsymbol{w}}_{\xi_j} + \tilde{\boldsymbol{w}}_{c_j}^{\mathrm{T}} \tilde{\boldsymbol{w}}_{c_j} + \tilde{\boldsymbol{w}}_{a_j}^{\mathrm{T}} \tilde{\boldsymbol{w}}_{a_j}}{2} \tag{49}$$

where $\tilde{\boldsymbol{w}}_{\xi_j} = \hat{\boldsymbol{w}}_{\xi_j} - \boldsymbol{w}_{\xi_j}^*, \tilde{\boldsymbol{w}}_{\star_j} = \hat{\boldsymbol{w}}_{\star_j} - \boldsymbol{w}_{J_j}^*, \star$ stands for c and a.

Based on (25),(37),(40)–(43), one can obtain

$$\dot{V} = \sum_{j=1}^{n} \left(-\upsilon_j z_j^2 - \frac{z_j \hat{\boldsymbol{w}}_{a_j}^{\mathrm{T}} \boldsymbol{S}_{J_j}}{2} + z_j \epsilon_{\xi_j} \right) + \varphi z_1 z_2 + \sum_{j=2}^{n} z_j z_{j+1} - \frac{\left(2\varphi^2 + 3\right) z_1^2}{4}$$

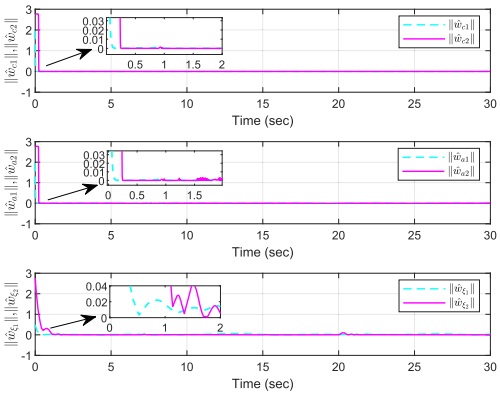


Fig. 4. Curves of the norm of the weight update laws.

$$-\sum_{j=2}^{n}\frac{7z_j^2}{4}+z_n s_2(u)-\sum_{j=1}^{n}\tilde{\boldsymbol{w}}_{a_j}^{\mathrm{T}}\boldsymbol{S}_{\mathcal{J}_j}\boldsymbol{S}_{\mathcal{J}_j}^{\mathrm{T}}\left(L_{a_j}\left(\hat{\boldsymbol{w}}_{a_j}-\hat{\boldsymbol{w}}_{c_j}\right)+L_{c_j}\hat{\boldsymbol{w}}_{c_j}\right)$$
$$-\sum_{j=1}^{n}\eta_j\tilde{\boldsymbol{w}}_{\xi_j}^{\mathrm{T}}\hat{\boldsymbol{w}}_{\xi_j}-\sum_{j=1}^{n}L_{c_j}\tilde{\boldsymbol{w}}_{c_j}^{\mathrm{T}}\boldsymbol{S}_{\mathcal{J}_j}\boldsymbol{S}_{\mathcal{J}_j}^{\mathrm{T}}\hat{\boldsymbol{w}}_{c_j}. \tag{50}$$

According to Young's inequality, we obtain

$$\varphi z_1 z_2\leq\frac{\varphi^2 z_1^2}{2}+\frac{z_2^2}{2}, \tag{51}$$
$$z_j z_{j+1}\leq\frac{z_j^2}{2}+\frac{z_{j+1}^2}{2}, \tag{52}$$
$$z_j\epsilon_{\xi_j}\leq\frac{z_j^2}{2}+\frac{\bar{\epsilon}_{\xi_j}^2}{2}, \tag{53}$$
$$-\frac{z_j\hat{\boldsymbol{w}}_{a_j}^{\mathrm{T}}\boldsymbol{S}_{\mathcal{J}_j}}{2}\leq\frac{z_j^2}{4}+\frac{\left(\hat{\boldsymbol{w}}_{a_j}^{\mathrm{T}}\boldsymbol{S}_{\mathcal{J}_j}\right)^2}{4}, \tag{54}$$
$$z_n s_2(u)\leq\frac{z_n^2}{2}+\frac{\bar{s}_2^2}{2}. \tag{55}$$

From the above discussion, one can obtain

$$\dot{V}\leq\sum_{j=1}^{n}\left(-v_j z_j^2-\eta_j\tilde{\boldsymbol{w}}_{\xi_j}^{\mathrm{T}}\hat{\boldsymbol{w}}_{\xi_j}-L_{c_j}\tilde{\boldsymbol{w}}_{c_j}^{\mathrm{T}}\boldsymbol{S}_{\mathcal{J}_j}\boldsymbol{S}_{\mathcal{J}_j}^{\mathrm{T}}\hat{\boldsymbol{w}}_{c_j}\right)+\sum_{j=1}^{n}\left(L_{a_j}-L_{c_j}\right)\tilde{\boldsymbol{w}}_{a_j}^{\mathrm{T}}\boldsymbol{S}_{\mathcal{J}_j}\boldsymbol{S}_{\mathcal{J}_j}^{\mathrm{T}}\hat{\boldsymbol{w}}_{c_j}$$
$$+\sum_{j=1}^{n}\frac{\bar{\epsilon}_{\xi_j}^2}{2}+\frac{\bar{s}_2^2}{2}-\sum_{j=1}^{n}L_{a_j}\tilde{\boldsymbol{w}}_{a_j}^{\mathrm{T}}\boldsymbol{S}_{\mathcal{J}_j}\boldsymbol{S}_{\mathcal{J}_j}^{\mathrm{T}}\hat{\boldsymbol{w}}_{a_j}+\sum_{j=1}^{n}\frac{\left(\hat{\boldsymbol{w}}_{a_j}^{\mathrm{T}}\boldsymbol{S}_{\mathcal{J}_j}\right)^2}{4}. \tag{56}$$

According to $\tilde{\boldsymbol{w}}_{\xi_j}=\hat{\boldsymbol{w}}_{\xi_j}-\boldsymbol{w}_{\xi_j}^*$ and $\tilde{\boldsymbol{w}}_{\star j}=\hat{\boldsymbol{w}}_{\star j}-\boldsymbol{w}_{\mathcal{J}_j}^*$, it follows that

$$\tilde{\boldsymbol{w}}_{\xi_j}^{\mathrm{T}}\hat{\boldsymbol{w}}_{\xi_j}=\frac{\tilde{\boldsymbol{w}}_{\xi_j}^{\mathrm{T}}\tilde{\boldsymbol{w}}_{\xi_j}+\hat{\boldsymbol{w}}_{\xi_j}^{\mathrm{T}}\hat{\boldsymbol{w}}_{\xi_j}-\boldsymbol{w}_{\xi_j}^{*\mathrm{T}}\boldsymbol{w}_{\xi_j}^*}{2}, \tag{57}$$
$$\tilde{\boldsymbol{w}}_{\star j}^{\mathrm{T}}\boldsymbol{S}_{\mathcal{J}_j}\boldsymbol{S}_{\mathcal{J}_j}^{\mathrm{T}}\hat{\boldsymbol{w}}_{\star j}=\frac{1}{2}\tilde{\boldsymbol{w}}_{\star j}^{\mathrm{T}}\boldsymbol{S}_{\mathcal{J}_j}\boldsymbol{S}_{\mathcal{J}_j}^{\mathrm{T}}\tilde{\boldsymbol{w}}_{\star j}+\frac{1}{2}\hat{\boldsymbol{w}}_{\star j}^{T}\boldsymbol{S}_{\mathcal{J}_j}\boldsymbol{S}_{\mathcal{J}_j}^{T}\hat{w}_{\star j}-\frac{1}{2}\boldsymbol{w}_{\mathcal{J}_j}^{*\mathrm{T}}\boldsymbol{S}_{\mathcal{J}_j}\boldsymbol{S}_{\mathcal{J}_j}^{\mathrm{T}}\boldsymbol{w}_{\mathcal{J}_j}^*. \tag{58}$$

Utilizing Young's inequality gives rise to the basic relationship

$$\tilde{\boldsymbol{w}}_{a_j}^{\mathrm{T}}\boldsymbol{S}_{\mathcal{J}_j}\boldsymbol{S}_{\mathcal{J}_j}^{\mathrm{T}}\hat{\boldsymbol{w}}_{c_j}\leq\frac{\left(\tilde{\boldsymbol{w}}_{a_j}^{\mathrm{T}}\boldsymbol{S}_{\mathcal{J}_j}\right)^2+\left(\hat{\boldsymbol{w}}_{c_j}^{\mathrm{T}}\boldsymbol{S}_{\mathcal{J}_j}\right)^2}{2}. \tag{59}$$

Let $\Delta_j=\bar{\epsilon}_{\xi_j}^2/2+\eta_j\boldsymbol{w}_{\xi_j}^{*\mathrm{T}}\boldsymbol{w}_{\xi_j}^*/2+\left(L_{c_j}+L_{a_j}\right)\left(\boldsymbol{w}_{\mathcal{J}_j}^{*\mathrm{T}}\boldsymbol{S}_{\mathcal{J}_j}\right)^2/2$, $L_{a_j}>L_{c_j}$, $L_{c_j}>L_{a_j}/2>1/4$, $\Lambda_{\boldsymbol{S}_{\mathcal{J}_j}}^{\min}$ be the minimal eigenvalue of $\boldsymbol{S}_{\mathcal{J}_j}\boldsymbol{S}_{\mathcal{J}_j}^{\mathrm{T}}$, and $\Lambda_{\boldsymbol{\Upsilon}_j^{-1}}^{\max}$ be the maximal eigenvalue of $\boldsymbol{\Upsilon}_j^{-1}$, one has

$$\dot{V}\leq\sum_{j=1}^{n}\left(-v_j z_j^2-\frac{\eta_j}{2}\tilde{\boldsymbol{w}}_{\xi_j}^{\mathrm{T}}\tilde{\boldsymbol{w}}_{\xi_j}-\frac{L_{c_j}}{2}\tilde{\boldsymbol{w}}_{c_j}^{\mathrm{T}}\boldsymbol{S}_{\mathcal{J}_j}\boldsymbol{S}_{\mathcal{J}_j}^{\mathrm{T}}\tilde{\boldsymbol{w}}_{c_j}-\frac{L_{c_j}}{2}\tilde{\boldsymbol{w}}_{a_j}^{\mathrm{T}}\boldsymbol{S}_{\mathcal{J}_j}\boldsymbol{S}_{\mathcal{J}_j}^{\mathrm{T}}\tilde{\boldsymbol{w}}_{a_j}+\Delta_j\right)$$

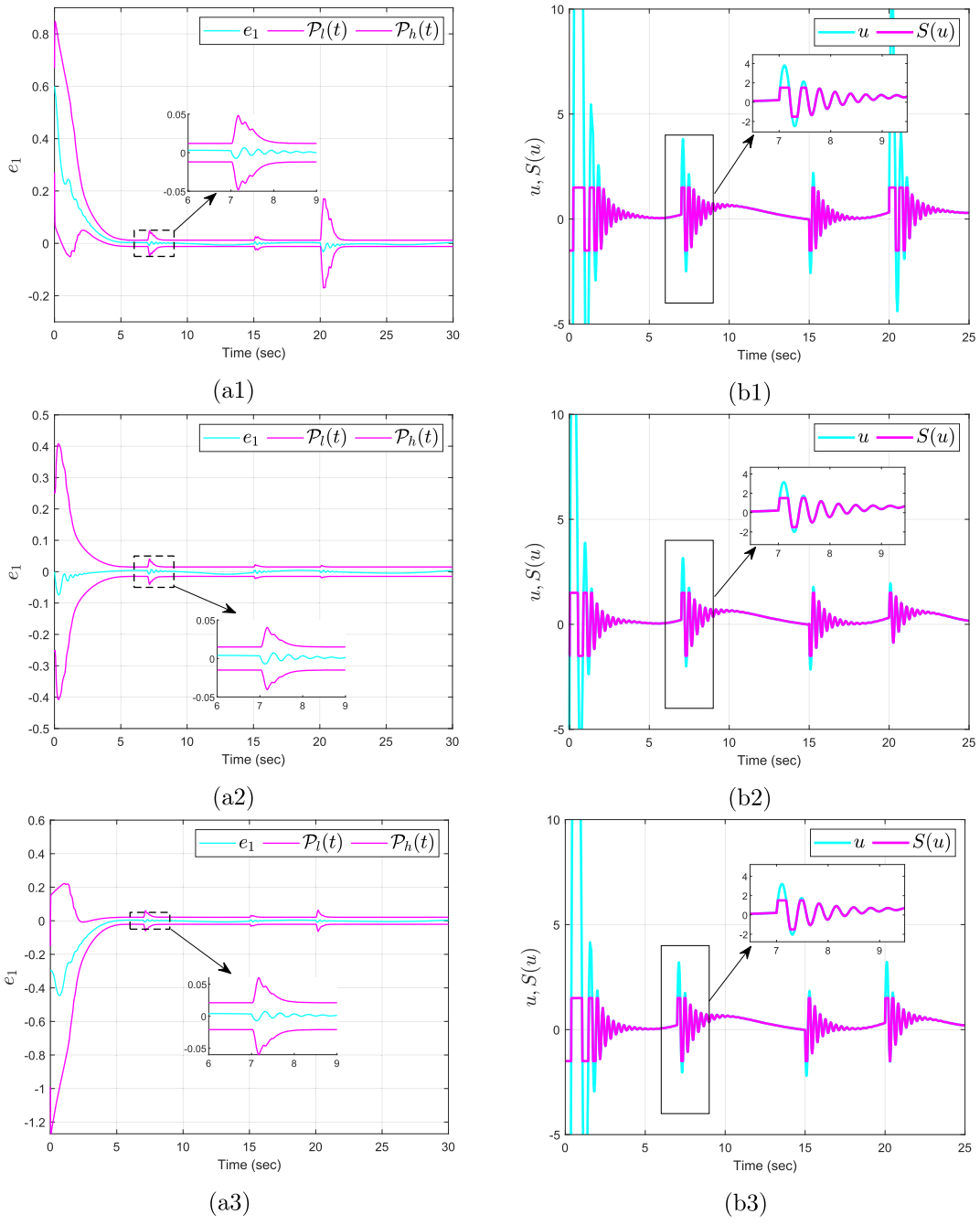


Fig. 5. (ai) and (bi) depict the tracking error trajectories and the associated control input signals under three distinct initial errors.

$$\leq \sum_{j=1}^{n}\left(-\upsilon_j z_j^2 - \frac{\eta_j \tilde{\boldsymbol{w}}_{\xi_j}^{\mathrm{T}} \boldsymbol{\Upsilon}_j^{-1} \tilde{\boldsymbol{w}}_{\xi_j}}{2\Lambda_{\boldsymbol{\Upsilon}_j^{-1}}^{\max}} - \frac{L_{c_j}}{2}\Lambda_{\boldsymbol{S}_{\mathcal{J}_j}}^{\min} \tilde{\boldsymbol{w}}_{c_j}^{\mathrm{T}} \tilde{\boldsymbol{w}}_{c_j} - \frac{L_{c_j}}{2}\Lambda_{\boldsymbol{S}_{\mathcal{J}_j}}^{\min} \tilde{\boldsymbol{w}}_{a_j}^{\mathrm{T}} \boldsymbol{S}_{\mathcal{J}_j} \boldsymbol{S}_{\mathcal{J}_j}^{\mathrm{T}} \tilde{\boldsymbol{w}}_{a_j} + \Delta_j\right). \tag{60}$$

Let $c = \min\{2\upsilon_j, \eta_j/\Lambda_{\boldsymbol{\Upsilon}_j^{-1}}^{\max}, L_{c_j}\Lambda_{\boldsymbol{S}_{\mathcal{J}_j}}^{\min}\}$, $\Delta = \sum_{j=1}^{n}\Delta_j$, we may consequently derive

$$\dot{V} \leq -cV + \Delta. \tag{61}$$

1) From (61), it is possible to deduce that

$$V \leq V(0) + \frac{\Delta}{c}. \tag{62}$$

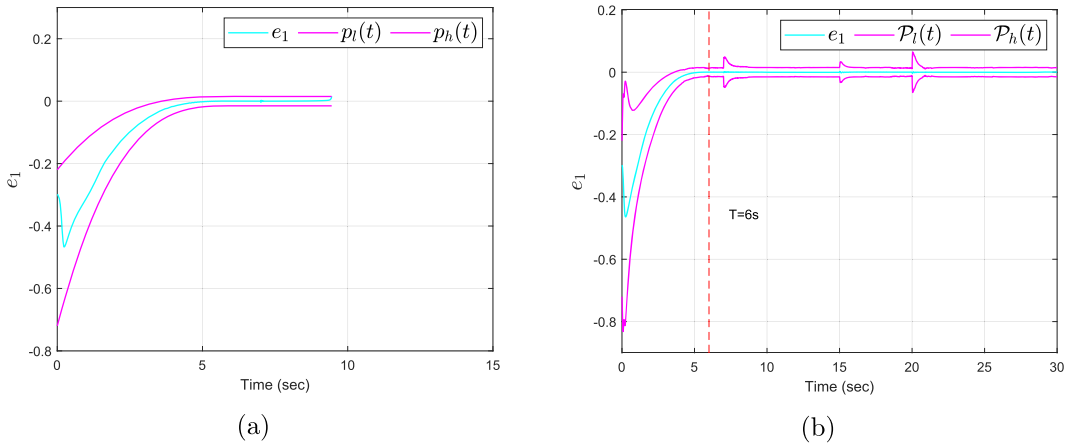


Fig. 6. (a) Tracking error curve under the proposed TyFPPC with $\mathbf{\Gamma}_1 = \mathbf{0}$; (b) Tracking error curve under the proposed TyFPPC with $\mathbf{\Gamma}_1 \neq \mathbf{0}$.

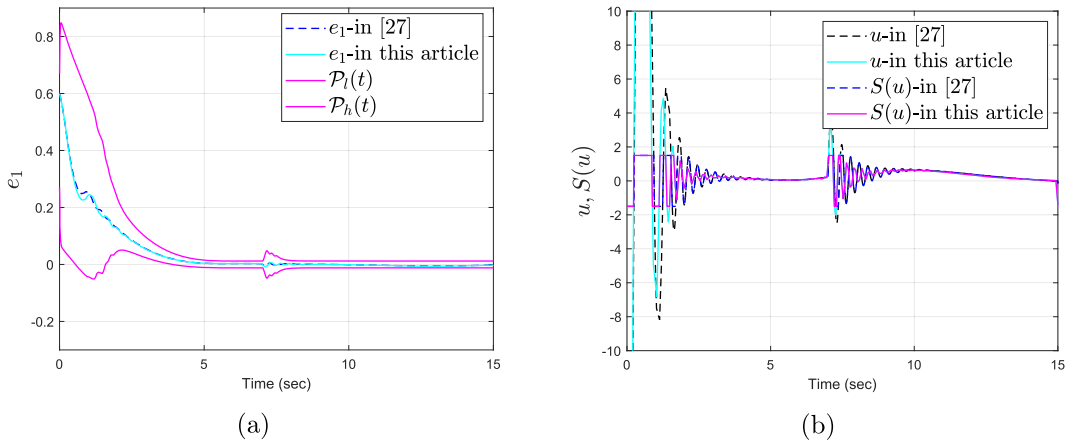


Fig. 7. (a) Trajectories of tracking errors obtained using the proposed approach versus that in [26]; (b) Trajectories of control signals obtained using the proposed approach versus that in [26].

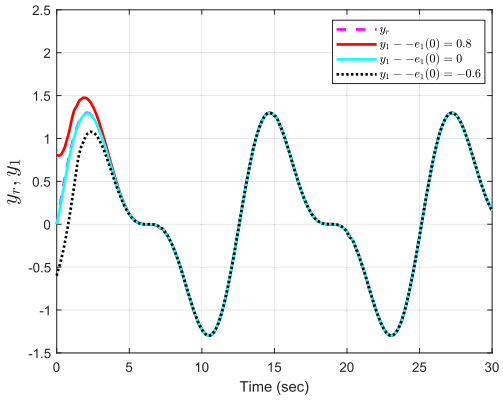


Fig. 8. Tracking trajectories under three distinct errors.

Based on the definition of V and (62), it follows that the variables z_j, $\tilde{\boldsymbol{w}}_{\xi_j}$, $\tilde{\boldsymbol{w}}_{c_j}$, and $\tilde{\boldsymbol{w}}_{a_j}$ remain bounded. On account of (12),(37),(40)–(43), one can obtain that $\hat{\alpha}_j^*$, $\hat{\boldsymbol{w}}_{\xi_j}$, $\hat{\boldsymbol{w}}_{c_j}$, $\hat{\boldsymbol{w}}_{a_j}$ remain bounded. As $z_1 = \varsigma, z_j = \chi_j - \hat{\alpha}_{j-1}^*$, thus χ_j and ς reamin bounded. As a result, all signals within the closed-loop system (1) reamin SGUUB.

2) It can be obtained from the definition of LF V and (62)

$$|z_1| = |\varsigma| \leq \sqrt{2\left(V_n(0) + \frac{\Delta}{c}\right)} \tag{63}$$

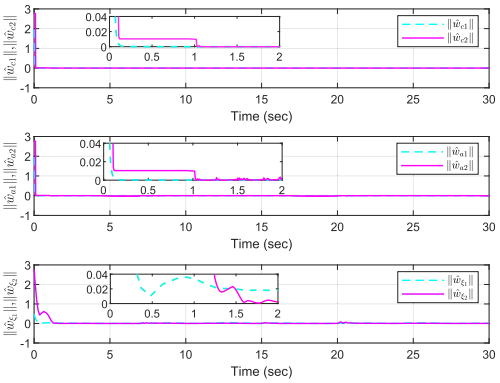


Fig. 9. Curves of the norm of the weight update laws.

implying that ς is bounded. According to the expression of $\delta(t)$ and $e_1(0) \in \left(\mathcal{P}_l(0), \mathcal{P}_h(0)\right)$, one can deduce that $\delta(0) \in (0,1)$. By combining the boundedness of ς and (10),one can further deduce that

$$\mathcal{P}_l(t) < e_1(t) < \mathcal{P}_h(t), \forall t \geq 0. \tag{64}$$

In summary, the system output can successfully track the desired signals while confining the tracking error within the TyFPPB. □

Remark 5. Note that the bounds on the approximation errors of the NNs and the supplementary system states affect the final convergence region. Based on the stability analysis, it can be concluded that the tracking accuracy can be improved by either increasing υ_j or reducing η_j. It is worth noting that the designed TyFPPB (7) also serves a crucial role, where ρ_T, and T respectively control the stability accuracy and convergence time, and the lower their values, the quicker the convergence rate and the greater the precision; b_1, b_2 in (9) control the speed of PPB expansion separately when input saturation occurs, and increasing b_2 can make the boundary contraction degree greater. In practice, the choice of parameters should also take into account specific application requirements.

4. Simulation results

Example 1. Let us consider the following ISNS

$$\begin{cases} \dot{\chi}_1 = \chi_2 + f_1(\chi_1) \\ \dot{\chi}_2 = S(u) + f_2(\bar{\chi}_2) \\ y = \chi_1 \end{cases} \tag{65}$$

where $f_1(\chi_1) = \sin \chi_1^2 \cos \chi_1$, $f_2(\bar{\chi}_2) = \chi_1(1+\chi_2^2)\sin \chi_2 - 1.2\chi_1\chi_2$. Select the design parameters as $\upsilon_1 = 150, \upsilon_2 = 3, L_{c_1} = 10, L_{c_2} = 180, L_{a_1} = 12, L_{a_2} = 350, \eta_1 = 75, \eta_2 = 50, \boldsymbol{\Upsilon}_1 = \boldsymbol{\Upsilon}_2 = 0.1\boldsymbol{I}_{24\times 24}$, $\rho_0 = 0.5, u_s = 1.5, T = 6, \rho_T = 0.03, l = 1, \gamma_1 = \gamma_2 = 0.2, b_1 = 5, b_2 = 0.5$, $y_r = \sin(0.5t)$. In Step 1, the critic, actor and identifier NNs are all set to contain 24 neurons, with their centers uniformly distributed within the interval $[-8, 8]$. $\boldsymbol{S}_{\star 1} = [s_{\star 1}^1, \ldots, s_{\star 1}^{24}]^{\mathrm{T}}$ ($\star$ stands for ξ, $\mathcal{J}$), $s_{\xi_1}^j = \exp((\chi_1 + 2j/3 - 8)^2/2^2), j = 1:24$. In Step 2, the critic, actor and identifier NNs are all set to contain 32 neurons, with their centers uniformly distributed within the interval $[-8, 8]$. $\boldsymbol{S}_{\star 2} = [s_{\star 2}^1, \ldots, s_{\star 2}^{32}]^{\mathrm{T}}$ with $s_{\xi_2}^j = \exp(||\bar{\chi}_2 - [8,8]^{\mathrm{T}} + 1/2[j,j]^{\mathrm{T}})||^2/2^2), j = 1:32$. $\hat{\boldsymbol{w}}_{\star_1}(0) = [0.4]_{24\times 1}$,$\hat{\boldsymbol{w}}_{\star_2}(0) = [0.4]_{32\times 1}$ with $\star = a, c, \xi$.

Fig. 3 illustrates the tracking trajectory corresponding to three distinct initial condition errors, and Fig. 5 illustrates the tracking error trajectories along with the associated control input signals for three distinct initial error conditions, verifying the system output successfully tracks the reference signals while maintaining the tracking errors within the TyFPPBs. Note that when $e_1(0) \neq 0$ (Subgraphs (a1) and (a3)), the proposed PPB presents a "tube" type, indicating that the proposed algorithm inherits the advantages of the EPPC methods. In addition, the proposed algorithm also further considered the intrinsic connection between input saturation and performance preset, i.e., when the control input exceeds the maximum threshold, the TyFPPB undergoes adaptive expansion, and once $|u| \leq u_s$, the TyFPPB adaptively reverts to the initial PPB, meaning that the proposed TyFPPC algorithm also achieves the balance between performance requirements and input security. Fig. 4 shows the curves of $\hat{\boldsymbol{w}}_{c_1}, \hat{\boldsymbol{w}}_{c_2}, \hat{\boldsymbol{w}}_{a_1}, \hat{\boldsymbol{w}}_{a_2}$ and $\hat{\boldsymbol{w}}_{\xi_1}, \hat{\boldsymbol{w}}_{\xi_2}$, respectively.

In order to validate the "flexibility" of the presented method, we present the simulation results with $\boldsymbol{\Gamma}_1 = \boldsymbol{0}$ and $\boldsymbol{\Gamma}_1 \neq \boldsymbol{0}$. In fact, without considering input saturation and optimization, the presented algorithm degenerates into the algorithm in [27]. As illustrated in Fig. 6 that under the input constraints, if $\boldsymbol{\Gamma}_1 = \boldsymbol{0}$, the proposed algorithm suffers from simulation interruption due to the singularity problem, whereas the proposed TyFPPC algorithm can still maintain stable control performance. The reason lies in that the proposed TyFPPC algorithm can achieve the balance between input security and performance preset by introducing a supplementary system, i.e., when input saturation occurs, singularity is avoided by relaxing the PPB.

Note that another advantage of the proposed method lies in the integration of the TyFPPB with the RL strategy, and the designed TyFPPB-RL optimal method can minimize the cost function of the control system while satisfying performance constraints and input constraints. To verify this, Fig. 7 presents a comparison between the approach in [26] and the proposed approach. With identical parameters and initial conditions, Fig. 7(a) and (b) demonstrate that the designed approach achieves reduced control effort and improved tracking accuracy compared to the approach in [26].

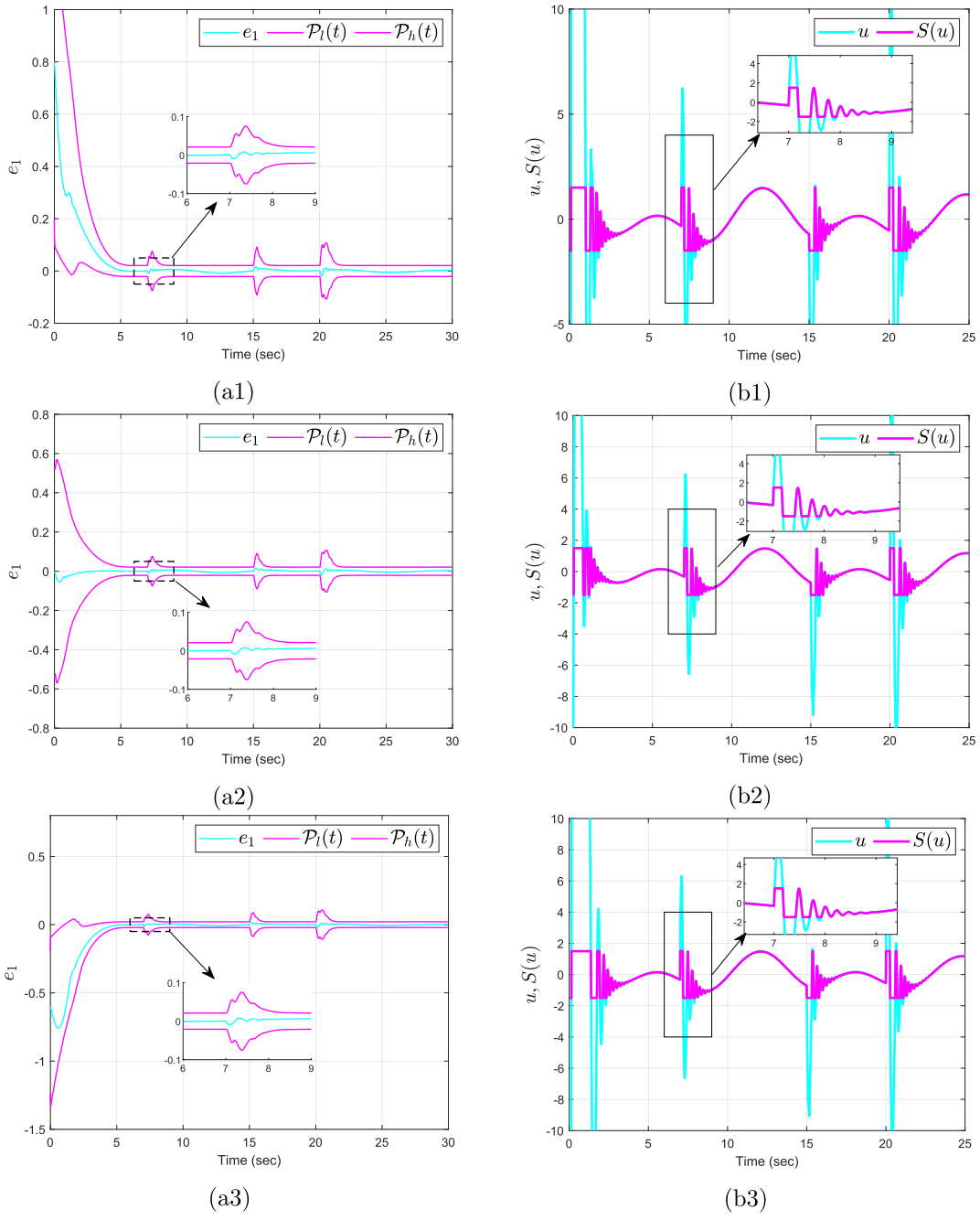


Fig. 10. (ai) and (bi) depict the tracking error trajectories and the associated control input signals under three distinct initial errors.

Example 2. A single-link robotic arm system [2] is considered below

$$\begin{cases} \dot{\chi}_1 = \chi_2 \\ \dot{\chi}_2 = \dfrac{S(u)}{M} - \dfrac{mgl\sin\chi_1}{2M} - \dfrac{B\chi_2}{M} \\ y = \chi_1 \end{cases} \tag{66}$$

where $g = 9.8m/s^2, m = 0.02kg, B = 1, l = 1m, M = 1kg$. Let $\upsilon_1 = 220, \upsilon_2 = 3.9, L_{c_1} = 10, L_{c_2} = 180, L_{a_1} = 12, L_{a_2} = 350, \eta_1 = 75, \eta_2 = 50, \Upsilon_1 = \Upsilon_2 = 0.1\boldsymbol{I}_{24\times 24}$, $u_s = 1.5, \rho_0 = 0.75, \rho_T = 0.03, T = 6, l = 1, \gamma_1 = \gamma_2 = 0.2, b_1 = 5, b_2 = 0.5$, $y_r = 0.5\sin(t) + \sin(0.5t)$. In Step 1, the critic, identifier and actor NNs are all designed to contain 24 neurons, with their centers uniformly distributed within the interval $[-8, 8]$. Regarding the basis function vector $\boldsymbol{S}_{\star 1} = [s^1_{\star 1}, \ldots, s^{24}_{\star 1}]^{\mathrm{T}}$ ($\star$ denotes ξ, $\mathcal{J}$) with $s^j_{\xi_1} = \exp((\chi_1 - 8 + 2j/3)^2/2^2), j = 1, \ldots, 24$. In

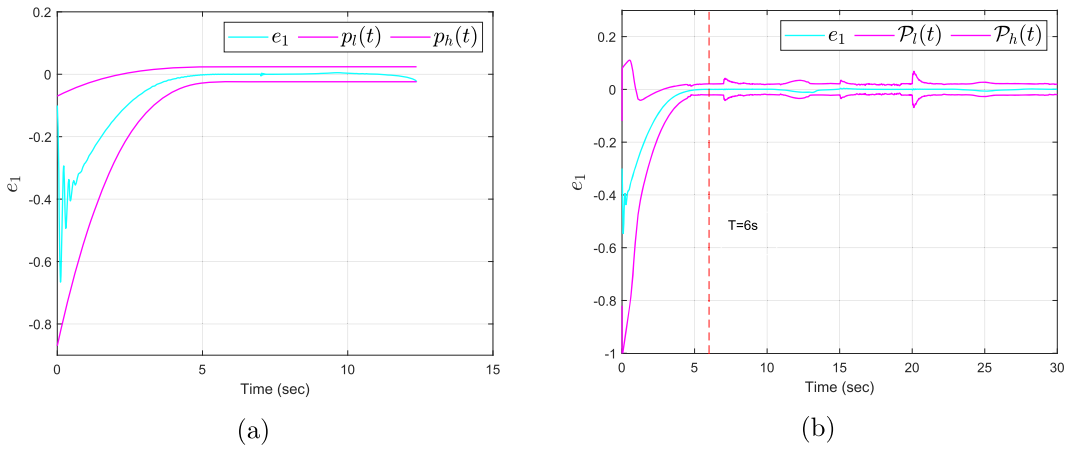


Fig. 11. (a) Tracking error curve under the proposed TyFPPC with $\mathbf{\Gamma}_1 = \mathbf{0}$; (b) Tracking error curve under the proposed TyFPPC with $\mathbf{\Gamma}_1 \neq \mathbf{0}$.

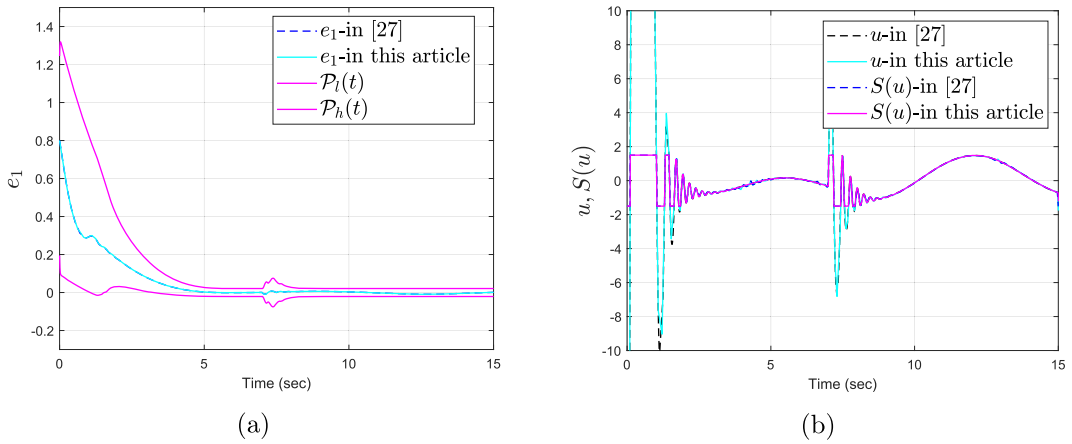


Fig. 12. (a) Trajectories of tracking errors obtained using the proposed approach versus that in [26]; (b) Trajectories of control signals obtained using the proposed approach versus that in [26].

Step 2, the critic, identifier and actor NNs are all designed to contain 32 neurons, with their centers uniformly distributed within the interval $[-8, 8]$. Regarding the basis function vector $\boldsymbol{S}_{\star 2} = [s^1_{\star 2}, \ldots, s^{32}_{\star 2}]^{\mathrm{T}}$ with $s^j_{\xi_2} = \exp(||\bar{\chi}_2 - [8,8]^{\mathrm{T}} + 1/2[j,j]^{\mathrm{T}})^2/2^2), j = 1, \ldots, 32$. $\hat{\boldsymbol{w}}_{\star_1}(0) = [0.4]_{24\times 1}, \hat{\boldsymbol{w}}_{\star_2}(0) = [0.4]_{32\times 1}$ with $\star = a, c, \xi$.

Fig. 8 illustrates the tracking trajectory corresponding to three distinct initial condition errors, and Fig. 10 illustrates the tracking error trajectories along with the associated control input signals for three distinct initial error conditions, verifying the system output successfully tracks the reference signals while maintaining the tracking errors within the TyFPPBs. Fig. 9 shows the curves of $\hat{\boldsymbol{w}}_{c_1}, \hat{\boldsymbol{w}}_{c_2}, \hat{\boldsymbol{w}}_{a_1}, \hat{\boldsymbol{w}}_{a_2}$ and $\hat{\boldsymbol{w}}_{\xi_1}, \hat{\boldsymbol{w}}_{\xi_2}$, respectively.

In order to verify the "flexibility" of the presented algorithm, we present the simulation results with $\mathbf{\Gamma}_1 = \mathbf{0}$ and $\mathbf{\Gamma}_1 \neq \mathbf{0}$. One can see from Fig. 11 that under the input constraints, if $\mathbf{\Gamma}_1 = \mathbf{0}$, the proposed algorithm suffers from simulation interruption due to the singularity problem, whereas the proposed TyFPPC algorithm can still maintain stable control performance. Fig. 12 presents a comparison between the proposed approach and that in [26]. With identical parameters and initial conditions, Fig. 12(a) and (b) demonstrate that the designed approach achieves reduced control effort and improved tracking accuracy compared to the approach in [26].

5. Conclusion

A TyFPPB-RL control approach for ISNSs is presented in this article, the significant advantage of the presented TyFPPC algorithm lies in its ability to comprehensively consider multiple indicators, and also achieves the balance between performance requirements and input security. In addition, the RL strategy with the ICAS is utilized to reduce the control system's cost function while achieving control performance. Future research will extend the TyFPPB-RL control to ISNSs with stochastic disturbances.

CRediT authorship contribution statement

Ziyi Liu: Writing – original draft; **Yangang Yao:** Writing – review & editing; **Yu Kang:** Writing – review & editing; **Yunbo Zhao:** Writing – review & editing; **Jieqing Tan:** Writing – review & editing; **Lichuan Gu:** Writing – review & editing; **Qiang Li:** Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

This work was supported by the Natural Science Foundation of Anhui Province under Grant 2508085MF165; in part by Anhui Provincial Department of Education Research Project under Grant 2024AH040082; in part by the Open Fund of State Key Laboratory of Tea Plant Biology and Utilization under Grant SKLTOF20240112; and in part by the National Natural Science Foundation of China (62103394, 62173317, 62033012 and 62172135).

Data availability

No data was used for the research described in the article.

References

[1] Li W, Krstic M. Mean-nonovershooting control of stochastic nonlinear systems. IEEE Trans Autom Control 2021;66(12):5756–71.
[2] Yao Y, Tan J, Wu J, Zhang X. A unified fuzzy control approach for stochastic high-order nonlinear systems with or without state constraints. IEEE Trans Fuzzy Syst 2022;30(10):4530–40.
[3] Liu Y, Wang Y, Feng Y, Wu Y. Neural network-based adaptive boundary control of a flexible riser with input deadzone and output constraint. IEEE Trans Cybern 2022;52(12):13120–8.
[4] Zhang Y, Xie L, Xie X, Sun Z, Zhang K. Fuzzy adaptive control for stochastic nonstrict feedback systems with multiple time-delays: a novel lyapunov-krasovskii method. IEEE Trans Fuzzy Syst 2024;32(6):3815–24.
[5] Zhao Z, Zhang J, Liu Z, Mu C, Hong K. Adaptive neural network control of an uncertain 2-dof helicopter with unknown backlash-like hysteresis and output constraints. IEEE Trans Neural Netw Learn Syst 2023;34(12):10018–27.
[6] Bechlioulis C, Rovithakis G. Robust adaptive control of feedback linearizable mimo nonlinear systems with prescribed performance. IEEE Trans Autom Control 2008;53(9):2090–9.
[7] Hua C, Zhang L, Guan X. Decentralized output feedback adaptive nn tracking control for time-delay stochastic nonlinear systems with prescribed performance. IEEE Trans Neural Netw Learn Syst 2015;26(11):2749–59.
[8] Liu Y, Ma H. Adaptive fuzzy tracking control of nonlinear switched stochastic systems with prescribed performance and unknown control directions. IEEE Trans Syst Man Cybern Syst 2020;50(2):590–9.
[9] Zhou Q, Li H, Wang L, Lu R. Prescribed performance observer-based adaptive fuzzy control for nonstrict-feedback stochastic nonlinear systems. IEEE Trans Syst Man Cybern Syst 2018;48(10):1747–58.
[10] Polyakov A, Efimov D, Perruquetti W. Finite-time and fixed-time stabilization: implicit lyapunov function approach. Automatica 2015;51:332–40.
[11] Zhang W, Zhong S, Jiang X. Finite-time annular domain stability and asynchronous h∞ control for stochastic switching markov jump systems. IEEE Trans Autom Control 2024;69(9):6277–84.
[12] Yao Y, Kang Y, Zhao Y, Li P, Tan J. Unified fuzzy control of high-order nonlinear systems with multi-type state constraints. IEEE Trans Cybern 2024;54(4):2525–35.
[13] Yu X, Wang G, Jia L, Zhang H. Event-triggered practical fixed-time containment control for stochastic multi-agent systems with input delay. IEEE Trans Fuzzy Syst 2024;32(5):2762–74.
[14] Zhu Y, Yao Y, Kang Y, Zhao Y, Tan J, Liang X. Event-based enhancing prescribed performance control for stochastic non-triangular structure nonlinear systems: a mtbfs-based approach. Nonlinear Dyn 2025;113:533–45.
[15] Yao Y, Tan J, Wu J, Zhang X, He L. Prescribed tracking error fixed-time control of stochastic nonlinear systems. Chaos Solit Fractals 2022;160:112288.
[16] Song Y, Wang Y, Holloway J, Krstic M. Time-varying feedback for regulation of normal-form nonlinear systems in prescribed finite time. Automatica 2017;83:243–52.
[17] Yao Y, Kang Y, Zhao Y, Li P, Tan J. A novel prescribed-time control approach of state-constrained high-order nonlinear systems. IEEE Trans Syst Man Cybern Syst 2024;54(5):2941–51.
[18] Yao Y, Kang Y, Zhao Y, Li P, Tan J. Prescribed-time output feedback control for cyber-physical systems under output constraints and malicious attacks. IEEE Trans Cybern 2024;54(11):6518–30.
[19] Xin B, Cheng S, Wang Q, Chen J, Deng F. Fixed-time prescribed performance consensus control for multiagent systems with nonaffine faults. IEEE Trans Fuzzy Syst 2023;31(10):3433–46.
[20] Liang Z, Wang Z, Zhao J, Wong P, Yang Z, Ding Z. Fixed-time and fault-tolerant path-following control for autonomous vehicles with unknown parameters subject to prescribed performance. IEEE Trans Syst Man Cybern Syst 2023;53(4):2363–73.
[21] Zhao K, Song Y, Ma T, He L. Prescribed performance control of uncertain Euler-Lagrange systems subject to full-state constraints. IEEE Trans Neural Netw Learn Syst 2018;29(8):3478–89.
[22] Liu Y, Zhang H, Wang Y, Ren H, Li Q. Adaptive fuzzy prescribed finite-time tracking control for nonlinear system with unknown control directions. IEEE Trans Fuzzy Syst 2022;30(6):1993–2003.
[23] Yao Y, Xu Z, Kang Y, Zhao Y, Tan J, Gu L, et al. Sliding flexible performance preset boundary-based fuzzy control for input saturated discrete-time nonlinear systems. IEEE Trans Autom Sci Eng 2026. https://doi.org/10.1109/TASE.2026.3675746
[24] Sui S, Chen C, Tong S. A novel adaptive nn prescribed performance control for stochastic nonlinear systems. IEEE Trans Neural Netw Learn Syst 2021;32(7):3196–205.
[25] Sun K, Guo R, Qiu J. Fuzzy adaptive switching control for stochastic systems with finite-time prescribed performance. IEEE Trans Cybern 2022;52(9):9922–30.
[26] Yao Y, Kang Y, Zhao Y, Li P, Tan J. Flexible prescribed performance output feedback control for nonlinear systems with input saturation. IEEE Trans Fuzzy Syst 2024;32(11):6012–22.
[27] Shi Y, Yi B, Xie W, Zhang W. Enhancing prescribed performance of tracking control using monotone tube boundaries. Automatica 2024;159:111304.

[28] Yao Y, Kang Y, Zhao Y, Li P, Tan J. Sliding flexible prescribed performance control for input saturated nonlinear systems. IEEE Trans Fuzzy Syst 2025;33(4):1241–52.
[29] Ji R, Ge SS. Event-triggered tunnel prescribed control for nonlinear systems. IEEE Trans Fuzzy Syst 2024;32(1):90–101.
[30] Yao Y, Kang Y, Zhao Y, Li P, Tan J. Dual flexible prescribed performance control of input saturated high-order nonlinear systems. IEEE Trans Cybern 2025;55(3):1147–58.
[31] Ren HR, Cao L, Ma H, Li HY. Dynamic event-triggered-based fuzzy adaptive pinning control for multiagent systems with output saturation. IEEE Trans Fuzzy Syst 2025;33(4):1277–86.
[32] Shen H, Yu X, Yan H, Park JH, Wang J. Robust fixed-time sliding mode attitude control for a 2-dof helicopter subject to input saturation and prescribed performance. IEEE Trans Transp Elect 2025;11(1):1223–33.
[33] Yong K, Chen M, Shi Y, Wu Q. Flexible performance-based robust control for a class of nonlinear systems with input saturation. Automatica 2020;122:109288.
[34] Ji R, Yang B, Ma J, Ge SS. Saturation-tolerant prescribed control for a class of mimo nonlinear systems. IEEE Trans Cybern 2022;52(12):13012–26.
[35] Ji R, Li D, Ge SS. Saturation-tolerant prescribed control for nonlinear time-delay systems. IEEE Trans Fuzzy Syst 2023;31(8):2495–508.
[36] Ji R, Ge SS, Li D. Saturation-tolerant prescribed control for nonlinear systems with unknown control directions and external disturbances. IEEE Trans Cybern 2024;54(2):877–89.
[37] Xie H, Zong G, Yang D, Zhao X, Yi Y. Flexible-fixed-time-performance-based adaptive asymptotic tracking control of switched nonlinear systems with input saturation. IEEE Trans Autom Sci Eng 2024;21(4):6371–82.
[38] Xu B, Shou Y, Wang X, Shi P. Finite-time composite learning control of strict-feedback nonlinear system using historical stack. IEEE Trans Cybern 2023;53(9):5777–87.
[39] Li H, Wu Y, Chen M. Adaptive fault-tolerant tracking control for discrete-time multiagent systems via reinforcement learning algorithm. IEEE Trans Cybern 2021;51(3):1163–74.
[40] Fernandez B, Graa M, Osa-Amilibia J, Larrucea X. Actor-critic continuous state reinforcement learning for windturbine control robust optimization. Inf Sci 2022;591:365–80.
[41] Wen G, Chen C. Optimized backstepping consensus control using reinforcement learning for a class of nonlinear strict-feedback dynamic multi-agent systems. IEEE Trans Neural Netw Learn Syst 2023;34(3):1524–36.
[42] Bai W, Li T, Long Y, Chen C, Xiao Y, Li W, et al. A novel adaptive control design for a class of nonstrict-feedback discrete-time systems via reinforcement learning. IEEE Trans Syst Man Cybern Syst 2024;54(2):1250–62.
[43] Bai W, Chen D, Zhao B, D'Ariano A. Reinforcement learning control for a class of discrete-time non-strict feedback multi-agent systems and application to multi-marine vehicles. IEEE Trans Intel Veh 2025;10(5):3613–25.
[44] Zhu H, Li Y, Tong S. Dynamic event-triggered reinforcement learning control of stochastic nonlinear systems. IEEE Trans Fuzzy Syst 2023;31(9):2917–28.
[45] Yan C, Xia J, Park J, Xie X. Reinforcement learning-based adaptive event-triggered fuzzy control for cyclic switched stochastic nonlinear systems with actuator faults. IEEE Trans Fuzzy Syst 2024;32(3):1131–43.
[46] Yao Y, Liu Z, Kang Y, Zhao Y, Tan J, Gu L, et al. Sliding flexible prescribed performance boundary-guided reinforcement learning control for input-constrained nonlinear systems. IEEE Trans Cybern 2026;56(2):968–79.
[47] Liu D, Mao Z, Jiang B, Yan X. Prescribed performance fault-tolerant control for synchronization of heterogeneous nonlinear mass using reinforcement learning. IEEE Trans Cybern 2024;54(9):5451–62.
[48] Liu X, Yan H, Zhou W, Wang N, Wang Y. Event-triggered optimal tracking control for underactuated surface vessels via neural reinforcement learning. IEEE Trans Ind Inf 2024;20(11):12837–47.