RL-Driven Self-Triggered Optimal Stabilization for Unknown Nonlinear Networked Industrial Systems: An Adaptive Fuzzy Identification Design

Jian Liu, *Senior Member, IEEE*, Jingjing Xia, Lijuan Zha, Jinliang Liu, *Senior Member, IEEE*, and Engang Tian, *Senior Member, IEEE*

***Abstract*—This article addresses the optimal stabilization of unknown nonlinear networked industrial systems (NISs) with limited communication resources, proposing an innovative self-triggered approximate optimal control framework fused with generalized fuzzy hyperbolic model (GFHM) and reinforcement learning (RL) gradient descent. To handle unknown dynamics without prior model information, a GFHM-based state identifier is constructed, leveraging its universal approximation, flexible structure, and fewer parameters to adaptively reconstruct nonlinear dynamics. A self-triggered mechanism predicts the next control update instant via software, eliminating continuous hardware monitoring and reducing online computing cost, network data transmission overhead, and actuator energy usage. Within the RL framework, a critic neural network (NN) is established, with weights tuned via gradient descent to approximate the Hamilton-Jacobi-Bellman (HJB) equation solution and derive the approximate optimal control policy. Theoretical analysis proves the closed-loop signals are uniformly ultimately bounded (UUB). Numerical simulations validate the scheme's effectiveness in ensuring optimal control performance while saving resources.**

***Index Terms*—Uniformly ultimately bounded (UUB), hamilton-jacobi-bellman (HJB) equation, neural network (NN), reinforcement learning (RL).**

Received 16 July 2026; accepted 4 August 2026. Date of publication 6 August 2026; date of current version 24 August 2026. This work was supported in part by the National Natural Science Foundation of China under Grant 62373252, Grant 62273174, and Grant 62001210, in part by the Research Project on Degree and Graduate Education of NUFE under Grant XYJS2202611, and in part by Young Elite Scientists Sponsorship Program by JSAST under Grant JSTJ-2025-452. *(Corresponding author: Jinliang Liu.)*

Jian Liu and Jingjing Xia are with the School of Computer and Artificial Intelligence, Nanjing University of Finance and Economics, Nanjing 210023, China (e-mail: liujian@nufe.edu.cn; 1120241383@stu.nufe.edu.cn).

Lijuan Zha is with the School of Science, Nanjing Forestry University, Nanjing 210037, China (e-mail: zhalijuan@njfu.edu.cn).

Jinliang Liu is with the School of Computer Science, Nanjing University of Information Science and Technology, Nanjing 210044, China (e-mail: 003768@nuist.edu.cn).

Engang Tian is with the School of Optical-Electrical and Computer Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China (e-mail: teg@usst.edu.cn).

Digital Object Identifier 10.1109/TICPS.2026.3721794

I. Introduction

With excellent flexible remote operation, nonlinear networked industrial systems (NISs) have gained widespread adoption across industrial automation, robotics, intelligent transportation and other industrial and engineering fields [1], [2], [3], [4]. However, real-world NISs have inherent nonlinear dynamics that are difficult to model accurately, hindering high-performance controller design [5]. Traditional Takagi-Sugeno (T-S) fuzzy identification suffer from exponentially increasing parameters, causing heavy computational overhead and high modeling costs. Recently, Li et al. modeled fractional-order nonlinear systems with the T-S fuzzy model, deriving the global model via fuzzy rule interpolation for controller design [6]. In Yuan's work [7], the T-S fuzzy model approximated turbine valve position nonlinearities, and the system was modeled via fuzzy rule interpolation and defuzzification. As highlighted in the study [8], the generalized fuzzy hyperbolic model (GFHM) holds prominent strengths including a flexible structure, low parameter count and universal approximation capability for smooth functions over compact sets, making it a more efficient alternative to system identification. To address the challenge brought by unknown nonlinear dynamics, this paper develops a state identifier following the core design logic of GFHM.

To realize optimal control for nonlinear NISs, the Hamilton-Jacobi-Bellman (HJB) equation must be solved, but traditional methods are limited by the curse of dimensionality and dependence on accurate models, leading to extreme difficulty in solving [9]. As elaborated in the reinforcement learning (RL) literature [10], [11], RL learns optimal control policies iteratively through system interaction without prior model information, and has become a mainstream tool for HJB equation solving [12], [13], [14]. In a study by Wen et al. [13], an adaptive dynamic programming (ADP)-boosted method could estimate the unknown nonlinearities for multiagent consensus control. Besides, to model unknown perturbations, Zhang et al. designed a neural network (NN) scheme for switched system optimal control. For optimal control, this paper adopts the RL framework, constructs a single-critic NN to approximate the HJB solution, and optimizes weights by gradient descent. RL and ADP both achieve optimal policy learning via environmental interaction [15], [16], so they are treated equivalently herein. The proposed method ensures single-critic NN weight convergence and enhances the HJB solution approximation accuracy.

Limited communication bandwidth poses another critical constraint on NISs [17], [18], [19]. Practically, the bandwidth of NISs becomes increasingly limited with the expansion of system scale and rising complexity of controlled objects [20]. Traditional time-triggered control updates control signals periodically, resulting in redundant data transmission and wasted energy [21], [22], [23]. For the sake of cutting resource waste and power loss, feedback control algorithms built upon RL or ADP have evolved from regular sampling to non-periodic transmission modes, departing from time-triggered architectures. In contrast to time-triggered frameworks, event-triggered control methods are responsive, releasing system signals to the communication network only when specific triggering conditions are met [24], [25], [26], [27]. This mechanism effectively lowers computing overhead, data transmission resource occupation and energy consumption, which has been confirmed in relevant studies. For instance, a periodic event-triggered mechanism was proposed for NIS under quantization and delay in [24].

Though event-triggered control reduces transmission by updating on state deviation, it requires continuous hardware state monitoring as noted in [28]. To address this, self-triggered control is introduced. It predicts the next update instant at the current one, removing the need for continuous hardware monitoring and allowing communication suspension on embedded devices [29], [30], [31], [32], [33], [34]. Specifically, researchers in [32] designed a self-triggered scheme with mask compensation for switched multi-agent systems. However, while self-triggered control strategies have received extensive investigations within model predictive control, distributed control and networked control, specific research on optimal control is quite rare [35], [36]. This paper integrates a self-triggered mechanism following related references. It predicts update moments from past triggering states to remove continuous hardware monitoring, enabling low-power communication on embedded devices and reducing bandwidth and energy consumption. A positive minimal bound on execution intervals is derived via theoretical analysis for avoiding Zeno behavior and ensure practicability.

Drawn from the above analysis, this work proposes an RL-driven self-triggered optimal control scheme with a GFHM-based state identifier for unknown nonlinear systems to address the hardware monitoring limitation of event-triggered control and reduce bandwidth usage. The core contributions are concluded in the following three aspects:

1) Unlike [6], [7] focusing on traditional fuzzy identification methods, we propose a GFHM-based state identifier. Specifically, the proposed method can adaptively approximate unknown nonlinear dynamics with fewer parameters and flexible structure without any prior system information, effectively reducing modeling complexity while guaranteeing approximation accuracy.
2) A resource-optimized self-triggered scheme (RoSTS) is embedded within the optimal control framework. Inspired by related research on self-triggered control, it predicts the next update instant through software relying on the state captured at the prior update moment, which eliminates the need for continuous hardware monitoring contrary to the event-triggered mechanism [25], [26], [27], saves communication bandwidth and energy consumption, and avoids Zeno behavior.
3) Compared with heuristic algorithms, e.g., particle swarm optimization that suffer from slow convergence [8], we design an optimal control framework built on RL gradient descent. Since the algorithm adopts a single-critic NN to estimate the solution to the HJB equation, it ensures rapid convergence and high approximation precision, and achieves near-optimal control performance for unknown nonlinear NISs.

The remainder of this work is structured as follows. Section II illustrates the system model. Section III details the design of the event-driven RNN identifier and presents NN-driven feedforward compensation. To ensure that all closed-loop signals satisfy the uniformly ultimately bounded (UUB), a Lyapunov stability analysis is conducted in Section IV. In Section V, we supply numerical cases to demonstrate the soundness of the presented approach. Lastly, Section VI provides concluding remarks.

II. System Model and Problem Description

Taking GFHM as the foundation, the system state identifier is set up in this paper. This section will first present the background information of GFHM, followed by the reconstruction of the considered unknown system.

A. Background of GFHM

First, the generalized fuzzy hyperbolic rule base is set forth in the following manner:

Definition 1: Taking a system with the input variable $\check{x} = [\check{x}_1, \check{x}_2, \ldots, \check{x}_p]^T \in \Re^p$ and the unique output $y \in \Re$ into account, a fuzzy rule base is considered a generalized fuzzy hyperbolic rule base if and only if it satisfies the following conditions.

1) The $\wp$th fuzzy rule, i.e., $\wp = 1, \ldots, 2^{\imath}, \imath = \sum_{i=1}^{p} n_i$, is given below:

IF $(\check{x}_1 - r_{11})$ is $\mathfrak{D}_{11}, \ldots, (\check{x}_1 - r_{1n_1})$ is $\mathfrak{D}_{1n_1}$ and $(\check{x}_2 - r_{21})$ is $\mathfrak{D}_{21}, \ldots, (\check{x}_2 - r_{2n_2})$ is $\mathfrak{D}_{2n_2}$ and $\ldots$ and $(\check{x}_p - r_{p1})$ is $\mathfrak{D}_{p1}, \ldots, (\check{x}_p - r_{pn_p})$ is $\mathfrak{D}_{pn_p}$;

THEN $y^{\wp} = E_{\mathfrak{D}_{11}} + \cdots + E_{\mathfrak{D}_{1n_1}} + E_{\mathfrak{D}_{21}} + \cdots + E_{\mathfrak{D}_{2n_2}} + \cdots + E_{\mathfrak{D}_{p1}} + \cdots + E_{\mathfrak{D}_{pn_p}}$, where $n_i (i = 1, 2, \ldots, p)$ stands for the quality of the transformation relative to $\check{x}_i$. Moreover, $r_{i\tau} \in \Re (\tau = 1, 2, \ldots, n_i)$ represents the transformation bias, and $\mathfrak{D}_{i\tau}$ is the fuzzy set pertaining to $(\check{x}_\tau - r_{i\tau})$, consisting of a positive subset $\mathbb{P}_{i\tau}^{+}$ and a negative subset $\mathbb{N}_{i\tau}^{-}$. Next, $\breve{x} = [\check{x}_1 - r_{11}, \ldots, \check{x}_p - r_{pn_p}]^T$ denotes the generalized input variable.

2) $E_{\mathfrak{D}_{i\tau}}$ in the "**THEN**" segment is dependent on $\mathfrak{D}_{i\tau}$ in the "**IF**" segment in terms of its presence. Specifically, $E_{\mathfrak{D}_{i\tau}}$ exists concomitantly with $\mathfrak{D}_{i\tau}$, and accordingly, $E_{\mathfrak{D}_{i\tau}}$ is non-existent when $\mathfrak{D}_{i\tau}$ is absent.

3) In view of all possible combinations of $\mathbb{P}_{i\tau}^{+}$ and $\mathbb{N}_{i\tau}^{-}$, the total number of fuzzy rules amounts to $2^{\imath}$.

To clarify the definition of GFHM and its universal approximation property, two lemmas are introduced below.

Lemma 1: With the p-input and single output system complying with Definition 1, the l-th component of $\breve{x}$ is expressed by $\breve{x}_l = \breve{x}_i - r_{i\tau}$ for $l = n_1 + \cdots + n_{i-1}$. The membership functions corresponding to $\breve{x}_l$ take the form of

$$\Phi_{\mathbb{P}_l^+}(\breve{x}_l) = e^{-\frac{1}{2}(\breve{x}_l - m_l)^2}, \Phi_{\mathbb{N}_l^-}(\breve{x}_l) = e^{-\frac{1}{2}(\breve{x}_l - m_l)^2}. \quad (1)$$

Therefore, the system can be obtained as

$$y = \theta + \mu^T tanh(\boldsymbol{M}\breve{x}) \quad (2)$$

where $\theta \in \Re$, $\mu = [\mu_1, \ldots, \mu_L]^T \in \Re$ and $\boldsymbol{M} = diag\{m_1, \ldots, m_L\} \in \Re^{L\times L}$ denote the bias, the weight vector and intrinsic weight, respectively. Besides, $tanh(\boldsymbol{M}\breve{x}) = [tanh(m_1\breve{x}_1), \ldots, tanh(m_L\breve{x}_L)]^T$. On this basis, (1) is regarded as the GFHM.

Lemma 2: Consider an unknown continuous function $\boldsymbol{K}(x)$. For any positive constant σ, one can find a fuzzy function $\kappa(x)$ from the GFHM set that meets the subsequent condition:

$$sup|\kappa(x) - \boldsymbol{K}(x)| < \sigma. \quad (3)$$

Remark 1: As detailed in Lemma 1, the definition of GFHM reveals its structure is analogous to that of conventional neural networks [37]. Utilizing advanced learning methodologies, the weight parameters of GFHM can be adaptively updated to realize the adaptive identification. For notational simplicity, the internal weight matrix $\boldsymbol{M}$ is designated as the identity matrix. Furthermore, Lemma 2 theoretically demonstrates the universal approximation property of GFHM in [38]. With these inherent properties, GFHM is suitable for approximating the unknown nonlinear system investigated herein, and the design of a GFHM-based system identifier is presented in the following subsection.

B. System Estimation Assisted by GFHM

Consider a continuous-time NIS with unmodeled dynamics in the following form:

$$\dot{x} = f(x) + \mathcal{G}u(t) \quad (4)$$

in which $x(t) \in \Re^p$ and $u(t) \in \Re^q$ correspond to the system state and the control input, respectively. The function $f(x) = [f_1(x), \ldots, f_p(x)]^T \in \Re^p$ signifies an unknown nonlinear mapping continuously differentiable on a compact set $\Omega \in \Re^p$. $\mathcal{G} = [\mathcal{G}_1, \ldots, \mathcal{G}_p]^T \in \Re^{p\times q}$ designates the input control matrix. Furthermore, we highlight that $f(x)$ and $\mathcal{G}$ are all unknown in our considered system, with only $x(t)$ and $u(t)$ both available for measurement.

Hence, every unknown $f_i(x)(i = 1, 2, \ldots, p)$ in (4) admits estiation via GFHM by virtue of Lemma 2. To estimate each $\dot{x}_i$, we employ a GFHM with the premise structure $\{n_1, \ldots, n_p\}$. To simplify our analysis, we represent all n_i as n in this section. This means we need p GFHMs to approximate the p-dimensional output system (4), and we express the GFHM-based estimation form of system as

$$\begin{cases} \dot{x}_1 = \mathscr{A}_1 + \mathscr{B}_1^T(\mathcal{X}_q - \mathscr{C}_1) + \epsilon_1 + \eta_1 x + \mathcal{G}_1 u \\ \quad\vdots \\ \dot{x}_p = \mathscr{A}_p + \mathscr{B}_p^T(\mathcal{X}_q - \mathscr{C}_p) + \epsilon_p + \eta_p x + \mathcal{G}_p u \end{cases} \quad (5)$$

with $\mathscr{A}_1, \ldots, \mathscr{A}_p \in \Re$. $\mathscr{B}_1, \ldots, \mathscr{B}_p \in \Re^{np}$ and $\mathscr{C}_1, \ldots, \mathscr{C}_p \in \Re^{np}$ denote the ideal weight matrics of the GFHM. $\mathcal{X}_q = x \otimes I_n \in \Re^{np}$, where I_n and $\otimes$ are the $n \times 1$ column vector of ones and the Kronecker product, respectively. The GFHM estimation errors $\epsilon_1, \ldots, \epsilon_p \in \Re$ are bounded. $\eta_1, \ldots, \eta_p$ respectively stand for the p entries of a preassigned stable design matrix.

We introduce an augmented operation in the following form:

$$\mathcal{X} = \Bigg[\underbrace{\mathcal{X}_q^T, \ldots, \mathcal{X}_q^T}_{p}\Bigg]^T \overset{\text{def}}{=} \Psi(\mathcal{X}_q) \in \Re^{np^2}.$$

Therefore, (5) can be established below:

$$\dot{x} = \bar{\mathcal{A}} + \bar{\mathcal{B}}^T tanh(\mathcal{X} - \bar{\mathcal{C}}) + \epsilon + \eta x + \mathcal{G}u \quad (6)$$

where $\bar{\mathcal{A}} = [\mathscr{A}_1, \ldots, \mathscr{A}_p]^T \in \Re^p$, $\bar{\mathcal{B}} = diag\{\mathscr{B}_1^T, \ldots, \mathscr{B}_p^T\}^T \in \Re^{np^2\times p}$, $\bar{\mathcal{C}} = [\mathscr{C}_1^T, \ldots, \mathscr{C}_p^T]^T \in \Re^{np^2}$ and the estimation error term $\epsilon = [\epsilon_1, \ldots, \epsilon_p]^T \in \Re^p$. In addition, $\eta = [\eta_1, \ldots, \eta_p]^T \in \Re^{p\times p}$ represents the stable matrix.

Accordingly, the dynamics of system (4) coincide with (6). Denoting the approximated state by $\hat{x}$, the state identification is given by

$$\dot{\hat{x}} = \hat{\mathcal{A}} + \hat{\mathcal{B}}^T tanh(\hat{\mathcal{X}} - \hat{\mathcal{C}}) + \eta\hat{x} + \hat{\mathcal{G}}u \quad (7)$$

where $\hat{\mathcal{A}}$, $\hat{\mathcal{B}}$, $\hat{\mathcal{C}}$ and $\hat{\mathcal{G}}$ represent the approximated values of $\bar{\mathcal{A}}$, $\bar{\mathcal{B}}$, $\bar{\mathcal{C}}$ and $\mathcal{G}$, respectively. Meanwhile, following the definition of $\mathcal{X}$, it holds that $\hat{\mathcal{X}} = \Psi(\hat{x} \otimes I_n)$.

Let the state identification error of the system be $\mathscr{E} = x - \hat{x}$, and the weight approximation errors be $\tilde{\mathcal{A}} = \bar{\mathcal{A}} - \hat{\mathcal{A}}$, $\tilde{\mathcal{B}} = \bar{\mathcal{B}} - \hat{\mathcal{B}}$, $\tilde{\mathcal{C}} = \bar{\mathcal{C}} - \hat{\mathcal{C}}$ and $\tilde{\mathcal{G}} = \mathcal{G} - \hat{\mathcal{G}}$. Noting that $\tilde{\mathcal{X}} = \mathcal{X} - \hat{\mathcal{X}}$, we have $\tilde{\mathcal{X}} = \Psi(\mathscr{E} \otimes I_n)$. By combining (6) and (7), the dynamic equation of the system identification error is derived as

$$\begin{aligned} \dot{\mathscr{E}} &= \tilde{\mathcal{A}} + \bar{\mathcal{B}}^T(\tanh(\mathcal{X} - \bar{\mathcal{C}}) - \tanh(\hat{\mathcal{X}} - \hat{\mathcal{C}})) \\ &\quad + \tilde{\mathcal{B}}^T \tanh(\hat{\mathcal{X}} - \hat{\mathcal{C}}) + \eta\mathscr{E} + \tilde{\mathcal{G}}u + \epsilon \end{aligned} \quad (8)$$

with the bounded error term ϵ.

C. Stability Analysis

Prior to presenting the adaptive weight update adjustments for the state identification (7), the subsequent assumption is indispensable for analyzing the stability of the state approximation error.

Assumption 1: For the existence of positive scalars $\varepsilon_1, \varepsilon_2$ and $\mathcal{B}_M$, it holds that $\tilde{\mathcal{C}}^T\tilde{\mathcal{C}} \leq \varepsilon_1\tilde{\mathcal{X}}^T\tilde{\mathcal{X}}$, $\tilde{\epsilon}^T\tilde{\epsilon} \leq \varepsilon_2\mathscr{E}^T\mathscr{E}$ and $||\bar{\mathcal{B}}|| \leq \mathcal{B}_M$.

In what follows, the update weight adjustments for the identifier (6) are specified in Theorem 1. Based on these, we establish the convergence properties of the weight approximation errors $\tilde{\mathcal{A}}$, $\tilde{\mathcal{B}}$, $\tilde{\mathcal{G}}$ and $\tilde{\mathcal{C}}$, as well as the system estimation errors $\mathscr{E}$.

Theorem 1: Under the condition of the unknown system dynamic (4), we implement the state identification (7) and formulate the tuning adjustments of $\hat{\mathcal{A}}$, $\hat{\mathcal{B}}$, $\hat{\mathcal{G}}$ and $\hat{\mathcal{C}}$ as

$$\dot{\hat{\mathcal{A}}} = \rho_1 \mathscr{E}, \qquad \dot{\hat{\mathcal{B}}} = \rho_{2,i}\mathscr{E}_i \tanh(\hat{\mathcal{X}}_q - \hat{\mathcal{C}}_i),$$
$$\dot{\hat{\mathcal{G}}} = \mathscr{E} u^T \pi, \qquad \dot{\hat{\mathcal{C}}} = -\rho_{3,i}(\mathscr{E} \otimes I_n) \tag{9}$$

with ρ_1, $\rho_{2,i}$ and $\rho_{3,i} > 0 (i = 1, \ldots, p)$ denoting learning rates, π being a symmetric positive definite matrix, and $\mathscr{E}_i$ the i-th element of $\mathscr{E}$. Consequently, the weight errors $\tilde{\mathcal{A}}$, $\tilde{\mathcal{B}}$, $\tilde{\mathcal{G}}$, $\tilde{\mathcal{C}}$ and the state approximation error $\mathscr{E}$ can achieve stability in the UUB sense with Assumption 1.

Proof: Take $\mathcal{L}_o = \mathcal{L}_1 + \mathcal{L}_2$ as the Lyapunov function, where the explicit forms of the sub-functions are

$$\mathcal{L}_1 = \frac{1}{2np}\tilde{\mathcal{X}}^T\tilde{\mathcal{X}} = \frac{1}{2}\mathscr{E}^T\mathscr{E} \tag{10}$$

and

$$\mathcal{L}_2 = \frac{1}{2}tr\left\{\tilde{\mathcal{A}}\rho_1^{-1}\tilde{\mathcal{A}}^T\right\} + \frac{1}{2}\sum_{i=1}^{p} tr\left\{\tilde{\mathcal{B}}_i\rho_{2,i}^{-1}\tilde{\mathcal{B}}_i^T\right\}$$
$$+ \frac{1}{2}tr\left\{\tilde{\mathcal{G}}\pi^{-1}\tilde{\mathcal{G}}^T\right\} + \frac{1}{2}\sum_{i=1}^{p} tr\left\{\tilde{\mathcal{C}}_i\rho_{3,i}^{-1}\tilde{\mathcal{C}}_i^T\right\}. \tag{11}$$

Inspired by (8), the derivative of $\mathcal{L}_1$ is inferred as

$$\begin{aligned}\dot{\mathcal{L}}_1 &= \mathscr{E}^T\dot{\mathscr{E}}\\ &= \mathscr{E}^T\tilde{\mathcal{A}} + \mathscr{E}^T\bar{\mathcal{B}}^T(\tanh(\mathcal{X} - \bar{\mathcal{C}}) - \tanh(\hat{\mathcal{X}} - \hat{\mathcal{C}}))\\ &\quad + \mathscr{E}^T\tilde{\mathcal{B}}^T\tanh(\hat{\mathcal{X}} - \hat{\mathcal{C}}) + \mathscr{E}^T\eta\mathscr{E} + \mathscr{E}^T\tilde{\mathcal{G}}u + \mathscr{E}^T\epsilon.\end{aligned}$$

From (10), we have $\tilde{\mathcal{X}}^T\tilde{\mathcal{X}} = np\mathscr{E}^T\mathscr{E}$. By virtue of the monotonicity of $tanh(\cdot)$, Assumption 1 and the Young inequality $2\mathtt{X}^T\mathtt{Y} \le \mathtt{X}^T\mathtt{X} + \mathtt{Y}^T\mathtt{Y}$, we arrive at the following expressions:

$$\begin{aligned}&\mathscr{E}^T\bar{\mathcal{B}}^T(\tanh(\mathcal{X} - \bar{\mathcal{C}}) - \tanh(\hat{\mathcal{X}} - \hat{\mathcal{C}}))\\ &\le \frac{1}{2}\mathscr{E}^T\bar{\mathcal{B}}^T\bar{\mathcal{B}}\mathscr{E} + \frac{1}{2}(\tanh(\mathcal{X} - \bar{\mathcal{C}}) - \tanh(\hat{\mathcal{X}} - \hat{\mathcal{C}}))^T\\ &\quad(tanh(\mathcal{X} - \bar{\mathcal{C}}) - \tanh(\hat{\mathcal{X}} - \hat{\mathcal{C}}))\\ &\le \frac{1}{2}\mathscr{E}^T\bar{\mathcal{B}}^T\bar{\mathcal{B}}\mathscr{E} + \frac{1}{2}(\tilde{\mathcal{X}} - \tilde{\mathcal{C}})^T(\tilde{\mathcal{X}} - \tilde{\mathcal{C}})\\ &\le \frac{1}{2}\mathscr{E}^T\bar{\mathcal{B}}^T\bar{\mathcal{B}}\mathscr{E} + \frac{(1+\varepsilon_1)np}{2}\mathscr{E}^T\mathscr{E} - \sum_{i=1}^{p}(\mathscr{E} \otimes I_n)^T\tilde{\mathcal{C}}_i\end{aligned} \tag{12}$$

and

$$\mathscr{E}^T\epsilon \le \frac{1}{2}\mathscr{E}^T\mathscr{E} + \frac{1}{2}\epsilon^T\epsilon \le \frac{1}{2}\mathscr{E}^T\mathscr{E} + \frac{1}{2}\varepsilon_2\mathscr{E}^T\mathscr{E}. \tag{13}$$

Utilizing (12) and (13), it is further derived as

$$\begin{aligned}\dot{\mathcal{L}}_1 &\le \mathscr{E}^T\tilde{\mathcal{A}} + \mathscr{E}^T\left[\frac{\mathcal{B}_M^2 + (1+\varepsilon_1)np + 1 + \varepsilon_2}{2}I_{p\times p} + \eta\right]\mathscr{E}\\ &\quad + \sum_{i=1}^{p}\mathscr{E}_i^T\mathcal{B}_i^T\tanh(\hat{\mathcal{X}}_q - \hat{\mathcal{C}}_i) + \mathscr{E}^T\tilde{\mathcal{G}}u\\ &\quad - \sum_{i=1}^{p}(\mathscr{E} \otimes I_n)^T\tilde{\mathcal{C}}_i.\end{aligned} \tag{14}$$

It can be easily deduced that $\dot{\tilde{\mathcal{A}}} = -\dot{\hat{\mathcal{A}}}$, $\dot{\tilde{\mathcal{B}}} = -\dot{\hat{\mathcal{B}}}$, $\dot{\tilde{\mathcal{G}}} = -\dot{\hat{\mathcal{G}}}$ and $\dot{\tilde{\mathcal{C}}} = -\dot{\hat{\mathcal{C}}}$, thus getting

$$\begin{aligned}\dot{\mathcal{L}}_2 &= -\mathscr{E}^T\tilde{\mathcal{A}} - \sum_{i=1}^{p}\mathscr{E}_i\tilde{\mathcal{B}}^T tanh^T(\hat{\mathcal{X}}_q - \hat{\mathcal{C}}_i)\\ &\quad - \mathscr{E}^T\tilde{\mathcal{G}}u + \sum_{i=1}^{p}(\mathscr{E} \otimes I_n)^T\tilde{\mathcal{C}}_i.\end{aligned} \tag{15}$$

By synthesizing the results from (14) and (15), we have the following formula

$$\begin{aligned}\dot{\mathcal{L}}_o &\le \mathscr{E}^T\left[\frac{\mathcal{B}_M^2 + (1+\varepsilon_1)np + 1 + \varepsilon_2}{2}I_{p\times p} + \eta\right]\mathscr{E}\\ &\overset{\text{def}}{=} \mathscr{E}^T\nu\mathscr{E}\end{aligned} \tag{16}$$

where $\nu = \frac{\mathcal{B}_M^2 + (1+\varepsilon_1)np + 1 + \varepsilon_2}{2}I_{p\times p} + \eta$. Based on the above derivation, we can conclude that choosing a suitable η can ensure the negative definiteness of δ, such that $\dot{\mathcal{L}}_o < 0$.

Consequently, it can be found that the GFHM-driven state identification (7) can precisely estimate the system (4). Moreover, upon completion of the full learning process, the approximated weights $\hat{\mathcal{A}}$, $\hat{\mathcal{B}}$, $\hat{\mathcal{G}}$ and $\hat{\mathcal{C}}$ can approach the neighborhood of $\bar{\mathcal{A}}$, $\bar{\mathcal{B}}$, $\mathcal{G}$ and $\bar{\mathcal{C}}$, with their final stable values denoted as $\mathcal{A}$, $\mathcal{B}$, $\mathcal{G}$ and $\mathcal{C}$, respectively; and we thereby acquire the reconstruction of the system

$$\dot{x} = \mathcal{A} + \mathcal{B}^T tanh(\mathcal{X} - \mathcal{C}) + \eta x + \mathcal{G}u \overset{\text{def}}{=} \hat{f}(x) + \mathcal{G}u \tag{17}$$

where $\hat{f}(x) = \mathcal{A} + \mathcal{B}^T tanh(\mathcal{X} - \mathcal{C}) + \eta x$. ∎

Remark 2: To streamline the following analysis, we maintain the notations $\dot{x}$, $f(x)$ and $\mathcal{G}$. Notably, we perform our upcoming derivation using this reconstructed system with $\mathcal{A}$, $\mathcal{B}$, $\mathcal{G}$ and $\mathcal{C}$ unchanged throughout.

III. Self-Triggered Optimal Control Design and Validity Analysis

In this section, we first derive the HJB equation by virtue of RoSTS. Then we formally prove that the system dynamics and the weight error dynamics of the critic NN are UUB when the expected control strategy is applied.

A. Optimal Control Framework

Define the perform function as

$$\mathscr{V}(x(t)) = \int_t^\infty \left(x^T(\tau)Qx(\tau) + u^T(\tau)Ru(\tau)\right)e^{\alpha(\tau - t)}d\tau$$

where α stands for the discount factor and $e^{\alpha(\tau-t)}$ is designed to ensure the convergence of $x(t)$. Since $\Xi(\Omega)$ denotes the admissible sets of $u(t)$, we conclude the optimal value corresponding to the performance function:

$$\mathscr{V}^*(x(t)) = \min_{u\in\Xi(\Omega)}\mathscr{V}\left(x(t)\right). \tag{18}$$

On the basis of the Bellman optimality principle, the optimal control problem in (18) can be used to resolve the following HJB equation

$$\underset{u\in\Xi(\Omega)}{min}\,\mathcal{H}\left(x(t),\nabla\mathscr{V}^*,u(x(t))\right)=0 \tag{19}$$

where $\mathcal{H}(\cdot)$ denotes the Hamiltonian function, with its detailed form expressed as:

$$\begin{aligned}&\mathcal{H}\left(x(t),u(x(t)),\nabla\mathscr{V}(x(t))\right)\\&=x^T(t)Qx(t)+u^T(x)Ru(x)-\alpha\mathscr{V}(x(t))\\&\quad+\left(\nabla\mathscr{V}(x(t))\right)^T\dot{x}(t)\end{aligned}$$

where $\nabla\mathscr{V}(x(t))$ denotes the partial derivative vector of $\mathscr{V}(x(t))$, that is $\nabla\mathscr{V}(x(t))=\frac{\partial\mathscr{V}}{\partial x}$.

According to Pontryagin's Minimum Principle, the following equation can be elicited:

$$\frac{\partial\mathcal{H}}{\partial u}=2Ru(x(t))+\mathcal{G}^T(x(t))\nabla\mathscr{V}^*(x(t)).$$

Therefore, we can simply acquire that

$$\begin{aligned}u^*(x(t))&=arg\underset{u\in\Xi(\Omega)}{min}\,\mathcal{H}\left(x(t),u(x(t)),\nabla\mathscr{V}^*(x(t))\right)\\&=-\frac{1}{2}R^{-1}\mathcal{G}^T(x(t))\nabla\mathscr{V}^*(x(t)).\end{aligned} \tag{20}$$

Recalling the optimal control input $u^*(x(t))$, we revise the HJB equation to integrate this input into its formulation, which is given by

$$\begin{aligned}&arg\underset{u\in\Xi(\Omega)}{min}\,\mathcal{H}\left(x(t),u(x(t)),\nabla\mathscr{V}^*(x(t))\right)\\&=x^T(t)Qx(t)+\left(u^*(x(t))\right)^T Ru^*(x(t))-\alpha\mathscr{V}(x(t))\\&\quad+(\nabla\mathscr{V}^*\left(x(t)\right))^T(f(x(t))+\mathcal{G}(x(t))u^*(t)).\end{aligned}$$

By incorporating RoSTS, the proposed optimal control strategy is updated exclusively at valid self-triggered update moments, that are satisfied, whose values are forecasted by software utilizing current system states and a specified functional relation. Hence, it is more readily implementable than event-triggered optimal control mechanism.

From the predictive results, the optimal control strategy gets renewed following the RoSTS, taking the form of $u_j=u\left(x(t_j)\right)$ at the forecasted triggering instant t_j. Meanwhile, consider $x(t_j)=x_j$ for convenience. As a result, our control objective lies in designing a self-triggered approximate optimal control strategy $u\left(x(t_j)\right)$ capable of ensuring UUB for all signals within the closed-loop system (4).

B. RL-Boosted Iteration Algorithm

Through the adoption of the NN approximation approach, we propose an RL-boosted iterative algorithm for solving the established HJB equation. To estimate the performance function $\mathscr{V}(x)$, we utilize the following single-layer NN:

$$\mathscr{V}^*(x)=\varpi^T\varphi(x)+\zeta(x) \tag{21}$$

where $\varpi\in\Re^{p_c}$ represents the designed optimal weight vector, with p_c being the number of hidden layer neurons. Meanwhile, $\varphi(x)=col\{\varphi_i\}(i=1,2,\ldots,p_c)$ denotes an undetermined activation function, in which each φ_i is independent of one another, $\zeta(x)$ is defined as the approximation error and $col(\cdot)$ refers to the column vector operator.

Then we derive its partial derivative written as

$$\nabla\mathscr{V}^*(x)=\nabla\varphi(x)^T\varpi+\nabla\zeta(x). \tag{22}$$

On account of (20), the optimal control strategy and the Hamiltonian are obtained given by

$$u^*(x)=u^*(x_j)=-\frac{1}{2}R^{-1}\mathcal{G}^T(x_j)\left(\nabla\varphi(x_j)^T\varpi+\nabla\zeta(x_j)\right)$$

and

$$\begin{aligned}\mathcal{H}\left(x,u^*(x_j),\varpi\right)&=x^TQx+(u^*(x_j))^TRu^*(x_j)+\delta\\&\quad-\alpha\varpi^T\varphi(x)+\varpi^T\nabla\varphi(x)\dot{x}=0\end{aligned} \tag{23}$$

where $\delta=-\alpha\zeta(x)+\nabla\zeta(x)\dot{x}$.

As a matter of fact, it is challenging to estimate the exact value of ϖ. Due to this unknown term, we use the following equation to derive its actual weight:

$$\hat{\mathscr{V}}(x)=\hat{\varpi}^T\varphi(x). \tag{24}$$

Subsequently, we compute the partial gradient of $\hat{\mathscr{V}}(x)$, expressed as

$$\nabla\hat{\mathscr{V}}(x)=\nabla\varphi(x)^T\hat{\varpi}. \tag{25}$$

Hence, the estimated optimal control input $\hat{u}(x)$ is given by

$$\hat{u}(x)=\hat{u}(x_j)=-\frac{1}{2}R^{-1}\mathcal{G}^T(x_j)\nabla\varphi(x_j)^T\hat{\varpi}. \tag{26}$$

The above strategy yields an estimated Hamiltonian, and its expression is as follows:

$$\begin{aligned}\hat{\mathcal{H}}\left(x,\hat{u}(x_j),\hat{\varpi}\right)&=x^TQx+\hat{u}^T(x_j)R\hat{u}(x_j)-\alpha\hat{\varpi}^T\varphi(x)\\&\quad+\hat{\varpi}^T\nabla\varphi(x)\dot{x}.\end{aligned} \tag{27}$$

Let the error between $\hat{\mathcal{H}}(\cdot)$ and $\mathcal{H}(\cdot)$ be defined as θ_z, leading to

$$\begin{aligned}\theta_z&=\hat{\mathcal{H}}\left(x,\hat{u}(x_j),\hat{\varpi}\right)-\mathcal{H}\left(x,u^*(x_j),\varpi\right)\\&=x^TQx+\hat{\varpi}^T\left(-\alpha\varphi(x)+\nabla\varphi(x)\dot{x}\right)+\hat{u}^T(x_j)R\hat{u}(x_j)\\&=x^TQx+\hat{u}^T(x_j)R\hat{u}(x_j)+\hat{\varpi}^T\Upsilon\end{aligned} \tag{28}$$

where $\Upsilon=-\alpha\varphi(x)+\nabla\varphi(x)\dot{x}$.

To effectively approximate $\mathscr{V}$ via $\hat{\mathscr{V}}$, we aim to realize the value of $u^*(x)$ as $\theta_z\to 0$. To accomplish this, we formally introduce an objective function:

$$\mathscr{E}=\frac{1}{2(1+\Upsilon^T\Upsilon)^2}\theta_z^T\theta_z. \tag{29}$$

Taking the gradient decent algorithm and the chain rule into account, we obtain

$$\dot{\hat{\varpi}} = -\rho_c \frac{\partial \mathscr{E}}{\partial \varpi} = -\rho_c \frac{\partial \mathscr{E}}{\partial \theta_z}\frac{\partial \theta_z}{\partial \varpi} = -\rho_c \frac{1}{(1+\Upsilon^T\Upsilon)^2}\frac{\partial \theta_z}{\partial \varpi}.$$

where ρ_c represents the learning rate that impacts algorithm convergence. By substituting the condition that $\frac{\partial \theta_z}{\partial \varpi} = \Upsilon - \hat{u}^T \mathcal{G}^T(x_j)\nabla\varphi^T(x_j)$ and the approximate optimality condition where the Hamiltonian's gradient with respect to $\hat{u}$ equals zero, we derive $\frac{\partial \theta_z}{\partial \varpi} = \Upsilon$. As a result, the update rule for $\hat{\varpi}$ is given by

$$\dot{\hat{\varpi}} = -\rho_c \frac{\Upsilon}{(1+\Upsilon^T\Upsilon)^2}\theta_z. \tag{30}$$

We next introduce $\tilde{\varpi} = \varpi - \hat{\varpi}$ to denote the weight approximation error, from which we deduce the following result that

$$\dot{\tilde{\varpi}} = -\rho_c \frac{\Upsilon\Upsilon^T}{(1+\Upsilon^T\Upsilon)^2}\tilde{\varpi} + \rho_c \frac{\Upsilon}{(1+\Upsilon^T\Upsilon)^2}\Theta_z \tag{31}$$

where $\Theta_z = -\nabla\varphi^T(x)\dot{x} + \alpha\varphi(x)$ indicates the residual error.

According to the updating law (30), $\hat{\varpi}$ can be computed at time instant t:

$$\begin{aligned}\hat{\varpi}(t) &= \hat{\varpi}(t-\Lambda_j)\\ &\quad - \rho_c\Lambda_j \frac{\Upsilon(t-\Lambda_j)}{(1+\Upsilon^T(t-\Lambda_j)\Upsilon(t-\Lambda_j))^2}\theta_z(t-\Lambda_j)\end{aligned}$$

where Λ_j indicates the intersampling time, thereby having the value of $\hat{u}(x_j)$. In other words, given the initial values of $\hat{\varpi}$ and $x(t)$, $\hat{u}(x_j)$ can be updated accordingly. Furthermore, the execution process of the optimal control algorithm based on NNs is illustrated within the context of Algorithm 1.

C. Convergence Verification

Prior to analyzing the convergence of the closed-loop system (4) under the proposed self-triggered approximation optimal control strategy, we first present the relevant assumptions.

Assumption 2: The drift dynamics $f(x)$ is Lipschitz continuous over the compact region Ω containing the origin. This ensures a unique solution $x(t)$ to the system (4) for $x_0 \in \Xi(\Omega)$ and the control input u. There exists a positive constant $\pounds_f$ meeting $||f(x)|| \leq \pounds_f||x||$, where $||\cdot||$ refers to the standard Euclidean norm, and the system (4) is stabilizable on Ω.

Assumption 3: It holds that the matrix $\mathcal{G}$ associated with the control input is uniformly norm-bounded by $0 \leq ||\mathcal{G}||_F \leq \pounds_g$ for any $x \in \Omega$, in which $\pounds_g > 0$ and $||\cdot||_F$ indicates the Frobenius norm.

Assumption 4: Suppose that the control input $u(x)$ is Lipschitz continuous over the compact set $\Xi(\Omega)$. Specifically, there exists a positive scalar $\pounds_u$ satisfying $||u(x(t)) - u(x(t_j))|| = ||u^* - u(x_j)|| \leq \pounds_u||h_k||$ with $h_j = x - x_j$.

Assumption 5: Norm bounds are imposed on $\nabla\varphi(x)$, $\nabla\zeta$, ζ and Θ_z, i.e., $||\nabla\varphi(x)|| \leq \varphi_{dM}$, $||\nabla\zeta|| \leq \zeta_{dM}$, $||\zeta|| \leq \zeta_M$ and $||\Theta_z|| \leq \Theta_{zM}$ with φ_{dM}, ζ_{dM}, ζ_M and Θ_{zM} being unknown positive scalars.

Assumption 6: Given the continuous differentiability of $\mathscr{V}^*(x)$ over its admissible set $\text{ß}(\Omega)$, one can verify that $\mathscr{V}^*(x)$ possesses an upper bound on $\text{ß}(\Omega)$, namely $||\mathscr{V}^*(x)|| < \mathscr{V}_M$.

Theorem 2: Given the unknown nonlinear system (4) and taking into account the Assumptions 2–6 and the proposed self-triggered estimated optimal control strategy (26), assume that positive scalars $\Gamma \in (0,1]$ and $\gamma \in (0,1)$ exist fulfilling

$$\Gamma(1-\gamma^2)x_j^T\mathcal{J}x_j \geq h_j^T\left[-\left(1-\frac{1}{\gamma^2}\right)\mathcal{J} + \mathcal{W}\right]h_j \tag{32}$$

where $\mathcal{J}$ and $\mathcal{W}$ denote positive definite diagonal matrices with suitable dimensions to be introduced subsequently. Under this condition, the UUB property for the closed-loop dynamics (4) is ensured.

Proof: Take the subsequent Lyapunov function for our stability analysis:

$$\mathfrak{L} = \mathfrak{L}_1 + \mathfrak{L}_2 \tag{33}$$

and the detailed information regarding $\mathfrak{L}$ is

$$\mathfrak{L}_1 = \frac{1}{2}x^Tx + \mathscr{V}^*(x), \quad \mathfrak{L}_2 = \frac{1}{2}\varpi^T\varpi.$$

First, the time derivative of $\mathfrak{L}_1$ is

$$\begin{aligned}\dot{\mathfrak{L}}_1 &= x^T\dot{x} + (\nabla\mathscr{V}^*(x))^T\dot{x}\\ &= (\nabla\mathscr{V}^*(x))^T f(x)\mathcal{G}\left(u^* + \hat{u}(x_j) - u^*\right)\\ &\quad + x^T\left(f(x) + \mathcal{G}\hat{u}(x_j)\right).\end{aligned} \tag{34}$$

According to (20) and (23), we have

$$\begin{aligned}&(\nabla\mathscr{V}^*(x))^T\mathcal{G} = -2(u^*)^TR\\ &(\nabla\mathscr{V}^*(x))^T\dot{x} = -x^TQx - (u^*)^TRu^* + \alpha\mathscr{V}^*(x).\end{aligned} \tag{35}$$

Combining Assumptions 2, 3 and 6, the above formulas (35), and the Young's inequality, it yields that

$$\begin{aligned}\dot{\mathfrak{L}}_1 &= x^T\left(f(x) + \mathcal{G}\hat{u}(x_j)\right) + (\nabla\mathscr{V}^*(x))^T\left(f(x) + \mathcal{G}u^*\right)\\ &\quad + 2(u^*)^TR\left(u^* - \hat{u}(x_j)\right)\\ &\leq x^T\pounds_f x + \pounds_g\left(\tfrac{1}{2}x^Tx + \tfrac{1}{2}\hat{u}^T(x_j)\hat{u}(x_j)\right) + \alpha\mathscr{V}_M\\ &\quad + 2(u^*)^TR\left(u^* - \hat{u}(x_j)\right) - x^TQx - (u^*)^TRu^*.\end{aligned} \tag{36}$$

Noting that $I_{p\times p}$ denoting the $p\times p$ dimensional identity matrix, it can be simply derived that

$$\begin{aligned}\dot{\mathfrak{L}}_1 &\leq -x^T\left[-\left(\pounds_f + \frac{1}{2}\pounds_g\right)I_{p\times p} + Q\right]x + \frac{1}{2}\pounds_g\hat{u}^T(x_j)\hat{u}(x_j)\\ &\quad + 2(u^*)^TR\left(u^* - \hat{u}(x_j)\right) - (u^*)^TRu^* + \alpha\mathscr{V}_M\\ &= -x^T\left[-\left(\pounds_f + \frac{1}{2}\pounds_g\right)I_{p\times p} + Q\right]x + \alpha\mathscr{V}_M\\ &\quad - (u^*)^T\left(R - \frac{1}{2}\pounds_g I_{q\times q}\right)u^*\end{aligned}$$

$$+ (u^*)^T \left(2R - \pounds_g I_{q\times q}\right) \left(u^* - \hat{u}(x_j)\right) + \frac{1}{2}\left(u^* - \hat{u}(x_j)\right)^T R \left(u^* - \hat{u}(x_j)\right). \tag{37}$$

According to (26) together with Assumption 4, the inequality can be acquired below:

$$\begin{aligned} ||u^* - \hat{u}(x_j)|| &\leq ||u^* - \hat{u}(x)|| + ||\hat{u}(x) - \hat{u}(x_j)|| \\ &\leq \frac{1}{2}R^{-1}\pounds_g\beta + \pounds_u||h_j|| \end{aligned} \tag{38}$$

where $\beta = \varphi_{dM}||\tilde{\varpi}|| + \zeta_{dM}$.

Therefore, (37) can be rewritten as

$$\begin{aligned} \dot{\mathfrak{L}}_1 \leq& -x^T\left[-\left(\pounds_f + \frac{1}{2}\pounds_g\right)I_{p\times p} + Q\right]x + \alpha\mathscr{V}_M \\ &+ \pounds_g\left(\frac{1}{4}R^{-2}\pounds_g^2\beta^2 + \pounds_u^2||h_j||^2 + R^{-1}\pounds_g\pounds_u\beta||h_j||\right) \\ &- (u^*)^T\left(R - \frac{1}{2}\pounds_g I_{q\times q}\right)u^* + x^T\mathcal{J}x - x^T\mathcal{J}x \\ &+ ||u^*||^T||2R - \pounds_g I_{q\times q}||_F||u^* - \hat{u}(x_j)|| \end{aligned} \tag{39}$$

In virtue of the Young's inequality, it is easy to obtain that

$$R^{-1}\pounds_g\pounds_u\beta||h_j|| \leq \frac{1}{2}R^{-2}\pounds_g^2\beta^2 + \frac{1}{2}\pounds_u^2||h_j||^2. \tag{40}$$

By adopting (40), the time difference of $\mathfrak{L}_1$ can be restated as

$$\begin{aligned} \dot{\mathfrak{L}}_1 \leq& -x^T\left[G - \left(\pounds_f + \frac{1}{2}\pounds_g\right)I_{p\times p} - \mathcal{J}\right]x - x^T\mathcal{J}x + \alpha\mathscr{V}_M \\ &+ \frac{3}{2}\pounds_g\pounds_u^2||h_j||^2 - (u^*)^T\left(R - \frac{1}{2}\pounds_g I_{q\times q}\right)u^* \\ &+ ||u^*||^T||2R - \pounds_g I_{q\times q}||_F||u^* - \hat{u}(x_j)|| + \frac{3}{4}R^{-2}\pounds_g^3\beta^2 \end{aligned} \tag{41}$$

where the last term based upon the Young's inequality yields the equivalent expression:

$$\begin{aligned} &||u^*||^T||2R - \pounds_g I_{q\times q}||_F||u^* - \hat{u}(x_j)|| \\ &\leq ||u^*||^T||2R - \pounds_g I_{q\times q}||_F\left(\frac{1}{2}R^{-1}\pounds_g\beta + \pounds_u||h_j||\right) \\ &\leq \frac{1}{4}\left((u^*)^T\pounds_g^2u^* + ||2R - \pounds_g I_{q\times q}||_F^2R^{-2}\beta^2\right) \\ &\quad + \frac{1}{2}(u^*)^T\pounds_u^2u^* + \frac{1}{2}||2R - \pounds_g I_{q\times q}||_F h_j^T h_j. \end{aligned} \tag{42}$$

Through incorporating (42) into (41) along with denoting $\breve{Q} = G - (\pounds_f + \frac{1}{2}\pounds_g)I_{p\times p} - \mathcal{J}$, we further get

$$\begin{aligned} \dot{\mathfrak{L}}_1 \leq& -x^T\breve{Q}x - x^T\mathcal{J}x + \alpha\mathscr{V}_M \\ &+ h_j^T\left(\frac{3}{2}\pounds_g\pounds_u^2 + \frac{1}{2}||2R - \pounds_g I_{q\times q}||_F^2\right)h_j \\ &- (u^*)^T\left[R - \frac{1}{2}\left(\frac{3}{2}\pounds_g^2 + \pounds_u^2\right)I_{q\times q}\right]u^* \\ &+ \left(\pounds_g^3 + ||2R - \pounds_g I_{q\times q}||_F^2\right)R^{-2}\beta^2. \end{aligned} \tag{43}$$

Recalling the expression that $\beta^2 = (\varphi_{dM}||\tilde{\varpi}|| + \zeta_{dM})^2 \leq 2\varphi_{dM}^2||\tilde{\varpi}||^2 + 2\zeta_{dM}^2$, it follows that

$$\begin{aligned} \dot{\mathfrak{L}}_1 \leq& -x^T\breve{Q}x - x^T\mathcal{J}x + h_j^T\mathcal{W}h_j - (u^*)^T\breve{R}u^* \\ &+ 2\xi\varphi_{dM}^2||\tilde{\varpi}||^2 + 2\xi\zeta_{dM}^2 + \alpha\mathscr{V}_M \end{aligned} \tag{44}$$

where $\xi = (\pounds_g^3 + ||2R - \pounds_g I_{q\times q}||_F^2)R^{-2}$, $\mathcal{W} = \frac{3}{2}\pounds_g\pounds_u^2 + \frac{1}{2}||2R - \pounds_g I_{q\times q}||_F^2$ and $\breve{R} = \pounds_g^3 + ||2R - \pounds_g I_{q\times q}||_F^2$ for the sake of convenience.

Since $h_j = x - x_j$, it is not hard to conclude that

$$\begin{aligned} x^T\mathcal{J}x &= (h_j + x_j)^T\mathcal{J}(h_j + x_j) \\ &= h_j^T\mathcal{J}h_j + x_j^T\mathcal{J}x_j + 2x_j^T\mathcal{J}h_j - \gamma^2x^T\mathcal{J}x \\ &\quad + \gamma^2x^T\mathcal{J}x - \frac{1}{\gamma^2}h_j^T\mathcal{J}h_j + \frac{1}{\gamma^2}h_j^T\mathcal{J}h_j \\ &\geq (1 - \gamma^2)x^T\mathcal{J}x + \left(1 + \frac{1}{\gamma^2}\right)h_j^T\mathcal{J}h_j. \end{aligned} \tag{45}$$

Under the condition of (32), one has

$$\begin{aligned} \dot{\mathfrak{L}}_1 \leq& -x^T\breve{Q}x - (1-\gamma^2)x^T\mathcal{J}x - \left(1 + \frac{1}{\gamma^2}\right)h_j^T\mathcal{J}h_j + \alpha\mathscr{V}_M \\ &+ h_j^T\mathcal{W}h_j - (u^*)^T\breve{R}u^* + 2\xi\varphi_{dM}^2||\tilde{\varpi}||^2 + 2\xi\zeta_{dM}^2 \\ \leq& -x^T\breve{Q}x - (1 - \gamma^2)x_j^T\mathcal{J}x_j + \alpha\mathscr{V}_M - (u^*)^T\breve{R}u^* \\ &+ \Gamma(1 - \gamma^2)x_j^T\mathcal{J}x_j + 2\xi\varphi_{dM}^2||\tilde{\varpi}||^2 + 2\xi\zeta_{dM}^2 \\ \leq& -\lambda_{min}(\breve{Q})||x||^2 + 2\xi\varphi_{dM}^2||\tilde{\varpi}||^2 + 2\xi\zeta_{dM}^2 + \alpha\mathscr{V}_M. \end{aligned} \tag{46}$$

Revoking (31), and therefore we get

$$\begin{aligned} \dot{\mathfrak{L}}_2 &= \tilde{\varpi}^T\dot{\tilde{\varpi}} \\ &= \tilde{\varpi}^T\left[-\rho_c\frac{\Upsilon\Upsilon^T}{(1 + \Upsilon^T\Upsilon)^2}\tilde{\varpi} + \rho_c\frac{\Upsilon}{(1 + \Upsilon^T\Upsilon)^2}\Theta_z\right]. \end{aligned} \tag{47}$$

Noticing that $(1 + \Upsilon^T\Upsilon) \geq 1$, one has

$$\begin{aligned} &\rho_c\tilde{\varpi}^T\frac{\Upsilon}{(1 + \Upsilon^T\Upsilon)^2}\Theta_z \\ &\quad\leq \frac{\rho_c}{2}\tilde{\varpi}^T\frac{\Upsilon\Upsilon^T}{(1 + \Upsilon^T\Upsilon)^2}\tilde{\varpi} + \frac{\rho_c}{2}\Theta_z^T\Theta_z. \end{aligned} \tag{48}$$

Defining $\Delta = \frac{\Upsilon\Upsilon^T}{(1+\Upsilon^T\Upsilon)^2}$ and substituting (48) and Assumption 4 into (47), it can be handled as

$$\begin{aligned} \dot{\mathfrak{L}}_2 &\leq -\frac{\rho_c}{2}\tilde{\varpi}^T\Delta\tilde{\varpi} + \frac{\rho_c}{2}\Theta_z^T\Theta_z \\ &\leq -\lambda_{min}(\Delta)||\tilde{\varpi}||^T + \frac{\rho_c}{2}\Theta_{zM}^2. \end{aligned} \tag{49}$$

Building on the foundation in (46) and (48), the following result can be acquired:

$$\begin{aligned} \dot{\mathfrak{L}} &= \dot{\mathfrak{L}}_1 + \dot{\mathfrak{L}}_2 \\ &\leq -\lambda_{min}(\breve{Q})||x||^2 - (\lambda_{min}(\Delta) - 2\xi\varphi_{dM}^2)||\tilde{\varpi}||^2 \end{aligned}$$

$$
\begin{aligned}
&+\frac{\rho_c}{2}\Theta_{zM}^2+2\xi\zeta_{dM}^2+\alpha\mathscr{V}_M\\
&=-\Psi_1||x||^2-\Psi_2||\tilde{\varpi}||^2+\Psi_3 \qquad (50)
\end{aligned}
$$

where $\Psi_1=\lambda_{min}(\breve{Q})$, $\Psi_2=\lambda_{min}(\Delta)-2\xi\varphi_{dM}^2$, $\Psi_3=\frac{\rho_c}{2}\Theta_{zM}^2+2\xi\zeta_{dM}^2+\alpha\mathscr{V}_M$.

In this regard, it is verified that $\dot{\mathfrak{L}}<0$ subject to Theorem 2. This necessitates the positive definiteness of both Ψ_1 and Ψ_2 and on the premise that at least one of the aforementioned inequalities holds:

$$
\begin{cases}||x||\geq\sqrt{\frac{\Psi_3}{\Psi_1}}\\ ||\tilde{\varpi}||\geq\sqrt{\frac{\Psi_3}{\Psi_2}}.\end{cases}
$$

∎

Remark 3: For Assumptions 2 and 3, practical systems are described by continuous nonlinear differential equations with a unique equilibrium. The drift function $f(x)$ and control input matrix $\mathcal{G}$ correspond to inertial, Coriolis and centripetal forces in robotic systems, and are thus norm-bounded and non-zero. Assumption 4 holds due to the property of infinite-horizon control, while Assumption 5 is guaranteed by finite practical approximation errors. All these assumptions are reasonable and applicable to RL-based control schemes in [8], [28].

IV. Self-Triggered Scheme Framework

This section presents the self-triggered scheme design for the RL-based algorithm, verifies its admissibility with Zeno behavior precluded, and analyzes the temporal and spatial computational complexity.

A. Self-Triggered Mechanism

Within the following part, the inter-execution interval $\Lambda_j=t_{j+1}-t_j(j=1,\ldots,\infty)$ takes the form $\Lambda_j=\mho(x_j)$, which has a positive lower bound, proving the RoSTS scheme valid. Before the proof, we define Zeno behavior as follows.

Definition 2: Set the nonlinear system Zeno on the condition that

$$
\lim_{j\to\infty} t_j=\sum_{j=0}^{\infty}\Lambda_j=t_\infty<\infty \qquad (51)
$$

where Λ_j represents the j-th sampling interval, and t_∞ is known as the Zeno time.

Theorem 3: The sequence of triggering times $\{t_j\}$ are constructed such that

$$
t_{j+1}=t_j+\frac{1}{\varrho}ln\left(\varrho\frac{\Gamma x_j^T\mathcal{J}x_j}{\jmath(x_j)}+1\right) \qquad (52)
$$

with $\varrho=2(1+\pounds_f^2)$ and $\jmath(x_j)=(2\pounds_f^2+\pounds_g^2\pounds_u^2)\lambda_{max}(\mathcal{W})x_j^Tx_j$. It follows that the self-triggering law (52) is well-posed for the nonlinear dynamics (4). In particular, the inter-execution time is expressed via

$$
\Lambda_j=\frac{1}{\varrho}ln\left(\varrho\frac{\Gamma x_j^T\mathcal{J}x_j}{\jmath(x_j)}+1\right)>0.
$$

Algorithm 1: Online RL Strategy Under RoSTS.

Initialization: Given the initial parameters $x(0)$ and $\hat{\varpi}(0)$. Set iteration index $k=0$ and the trigger iteration index $k_{tri}=0$.
Designate the initial system state $x(0)$ as the initial trigger state $x_0(0)$, and predict the subsequent trigger iteration index according to k_{tri} and $x_0(0)$.
while $k\leq k_{max}$ **do**
    if $k=k_{tri}$ **then**
        Predict the next trigger iteration index k through the current k_{tri} and the current triggered system state;
        Let $k\leftarrow k+1$;
    else
        Let $k\leftarrow k_{tri}$;
    end
end

Proof: Consider the following Lyapunov function candidate when $t\in[t_j,t_{j+1})$:

$$
\text{Ł}=h_j^T\mathcal{W}h_j. \qquad (53)
$$

Using the fact that $h_j=x-x_j$ together with the Young's inequality and Assumptions 2–3, computing the time derivative results in

$$
\begin{aligned}
\dot{\text{Ł}}&=2\mathcal{W}\left(h_j^Tf(x)+h_j^T\mathcal{G}\hat{u}(x_j)\right)\\
&\leq\mathcal{W}\left(h_j^Th_j+f(x)^Tf(x)\right)+\mathcal{W}\left[h_j^Th_j+\left(\mathcal{G}\hat{u}(x_j)\right)^T\mathcal{G}\hat{u}(x_j)\right]\\
&\leq 2h_j^T\mathcal{W}h_j+\pounds_f^2x^T\mathcal{W}x+\pounds_g^2\pounds_u^2x_j^T\mathcal{W}x_j \qquad (54)
\end{aligned}
$$

where it is not hard to derive

$$
\begin{aligned}
\pounds_f^2x^T\mathcal{W}x&=\pounds_f^2(h_j+x_j)^T\mathcal{W}(h_j+x_j)\\
&\leq 2\pounds_f^2h_j^T\mathcal{W}h_j+2\pounds_f^2x_j^T\mathcal{W}x_j. \qquad (55)
\end{aligned}
$$

Inserting (55) into (54), it yields

$$
\begin{aligned}
\dot{\text{Ł}}&\leq 2h_j^T\mathcal{W}h_j+2\pounds_f^2h_j^T\mathcal{W}h_j+2\pounds_f^2x_j^T\mathcal{W}x_j+\pounds_g^2\pounds_u^2x_j^T\mathcal{W}x_j\\
&=\left(2+2\pounds_f^2\right)h_j^T\mathcal{W}h_j+\left(2\pounds_f^2+\pounds_g^2\pounds_u^2\right)x_j^T\mathcal{W}x_j\\
&=\varrho\text{Ł}+\jmath(x_j). \qquad (56)
\end{aligned}
$$

With the aid of the comparison principle and the self-triggered law (52), when $t\in[t_j,t_{j+1})$, it follows that

$$
\dot{\text{Ł}}\leq\frac{\jmath(x_j)}{\varrho}\left(e^{\varrho(t-t_j)}-1\right)\leq\Gamma x_j^T\mathcal{J}x_j, \quad . \qquad (57)
$$

From (57), exponential stability of system (4) ensures $\Lambda_j>0$. This signifies that the inter-execution time is positive, avoiding Zeno behavior and bringing the proof to an end.

B. Computational Complexity

Following [28], we analysis the minimal computational cost of the proposed self-triggered optimal control. The Veronese map $\mathcal{M}$ transforms $x = [x_1, \ldots, x_p]^T \in \Re^p$ to $x^T x \in \Re^{p\times p}$, satisfying

$$\mathcal{M} = \left[x_1^2, x_1x_2, \ldots, x_1x_m, x_2^2, x_2x_3, \ldots, x_p^2\right]^T.$$

Hence, the space complexity is $\mathcal{O}(t_d)[p(p+1)/2]$, where t_d indicates the discretization step, $\mathcal{O}(t_d) = \lfloor \bar{\Lambda}_j/t_d \rfloor - \lfloor \underline{\Lambda}_j/t_d \rfloor$, $\bar{\Lambda}_j$ and $\underline{\Lambda}_j$ denote the maximum open-loop and minimum one-step computation time. Regarding time complexity, let t_f be the time single-instruction execution and the implementation necessitates $\mathcal{O}(t_d)[p(p+1)/2]t_f$ multiplications and additions, along with $\mathcal{O}(t_d)t_f$ during the operation phase.

Next, we focus on our key research concern. Observing that the candidate Lyapunov function (33) comprises x and $\mathscr{V}^*$, i.e., x and u from the starting instant to the infinite time, the proposed self-triggered control strategy (26) exhibits a space complexity of

$$\mathcal{N}^o = \mathcal{O}(t_d)\left[\frac{p(p+1)}{2} + \frac{q(q+1)}{2} + 1\right]$$

as it needs $\mathcal{O}(t_d)[p(p+1)/2]$ storage for calculating $x^T x$, $\mathcal{O}(t_d)[q(q+1)/2]$ for $u^T R u$ and $\mathcal{O}(t_d)$ for storing the integral at t_j. Simultaneously, its time complexity is

$$\mathcal{N}^t = [2p(p+1) + q(q+1) + \mathcal{O}(t_d)]t_f$$

given that the preprocessing time involves $\mathcal{O}(t_d)p(p+1)t_f$ for both $x^T x$ and $x^T Q x$, $\mathcal{O}(t_d)q(q+1)t_f$ for $u^T Q u$ within the integral, plus $\mathcal{O}(t_d)t_f$ for checking the inequality $\dot{\mathfrak{L}} \leq -\Psi_1||x||^2 - \Psi_2||\tilde{\varpi}||^2 + \Psi_3$ in the operation phrase. It is noteworthy that the minimum inter-execution time $\underline{\Lambda}_j$ relies on $\mathcal{N}^t$, while the maximum inter-execution time $\bar{\Lambda}_j$ is restricted by the upper bound of $\dot{\mathfrak{L}}$.

V. Numerical Results

Consider a rotational industrial mechanical system modeled as a torsional pendulum bar:

$$\begin{cases} \frac{d\vartheta}{dt} = \omega \\ \mathscr{R}\frac{d\omega}{dt} = u - \boldsymbol{m}\mathscr{G}\mathscr{L}\sin\vartheta - \mathscr{F}\frac{d\vartheta}{dt} \end{cases} \tag{58}$$

where $\boldsymbol{m} = 1/3kg$ and $\mathscr{L} = 2/3m$ correspond to the mass and length of the pendulum bar, respectively. We define the rotary inertia as $\mathscr{R} = 4/3\boldsymbol{m}\mathscr{L}^2$, the frictional factor as $\mathscr{F} = 0.2$ and the gravitational acceleration as $\mathscr{G} = 9.8\,m/s^2$. By substituting the system states ϑ and ω with x_1 and x_2, the system can be written in the state-space representation as

$$\dot{x} = \begin{bmatrix} x_2 \\ -\frac{\boldsymbol{m}\mathscr{G}\mathscr{L}}{\mathscr{R}}x_1 - \frac{\mathscr{F}}{\mathscr{R}}x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ \frac{1}{\mathscr{R}} \end{bmatrix} u. \tag{59}$$

Firstly, a GFHM state identifier with $n = 2$ and $\eta = -3I_{2\times2}$ is designed for system estimation. Set learning rates as $\rho_1 = 1$, $\rho_{2,1} = 2$, $\rho_{2,2} = 1.2$, $\mathscr{E}_2 = 1$, $\rho_{3,1} = 2$, with the initial state $\hat{x}_0 = [0.8, -0.9]^T$. Within the cost function, Q and R denote identity matrices of suitable sizes. The initial system state is $x_0 = [-1, 1]^T$, the critic NN activation $\varphi(x) = [x_1^2, x_1x_2, x_2^2]^T$ and initial weights $\hat{\varpi}_0 = [0.220, 0.309, 0.017]^T$. Weight learning rate $\rho_c = 0.4$. For self-triggering, $\varrho = 3.25$ and $\Gamma =$ 0.1, 0.3, 0.6, 0.8, 1 are chosen via trial and error to evaluate sampling sensitivity.

TABLE I

Sampling Counts Under Various Parameter Γ Values

Γ	0.1	0.3	0.6	0.8	1
Sampling counts	782	591	545	501	397
$\bar{\Lambda}_j$	0.363	0.617	0.849	0.959	0.981
$\underline{\Lambda}_j$	0.001	0.002	0.002	0.002	0.004
Convergence time	18.2	17.5	16.9	16.0	16.0

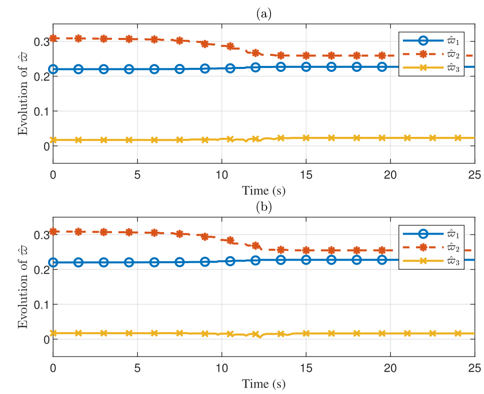

Fig. 1. Convergence process of the critic weights.

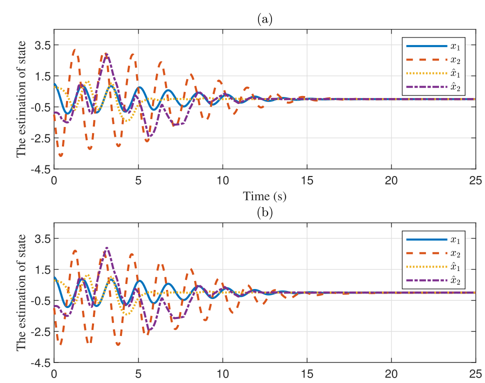

Fig. 2. Comparisons of the estimated states.

Using $\Gamma = 0.8$ in (a) and $\Gamma = 0.3$ in (b), simulations are reported within Figs. 1–5 and Table I. From Fig. 1, the critic NN weights converge to $[0.227, 0.260, 0.023]^T$. Fig. 2 depicts the states driven to zero rapidly and Fig. 3 shows the GFHM-based error stabilized around 17 s, verifying dynamics reconstruction. Fig. 4 compares the control inputs under the time-triggered mechanism (TTM) and the designed RoSTS, with the later piecewise constant and updating only at t_j. Fig. 5 shows the triggering interval and avoids Zeno behavior. Fig. 6 indicates sample count drops from 1001 to an average of 550, cutting about 45.1% frequency. Meanwhile, we can draw a conclusion from Table I that Γ increase cuts sample count and settling time to equilibrium. This means Γ can be selected based on transient response and resource cost demands.

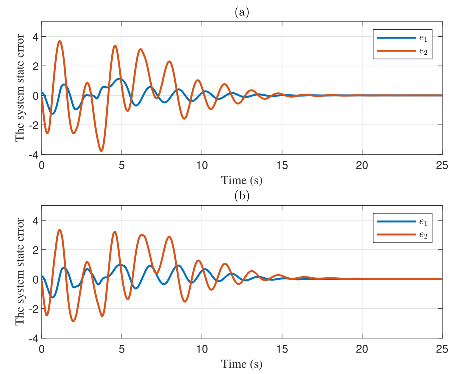

Fig. 3. Comparisons of the system state error.

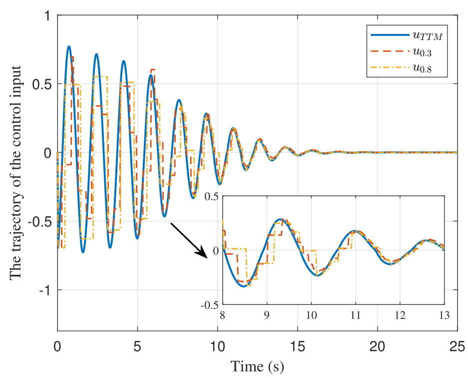

Fig. 4. Control inputs under different triggered mechanisms.

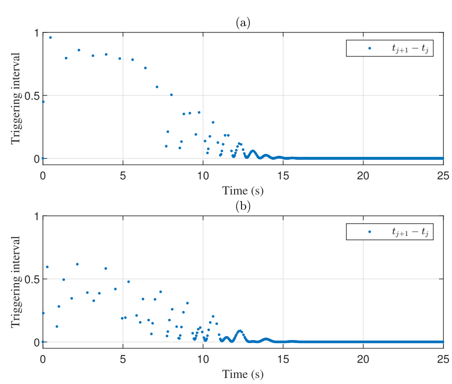

Fig. 5. Comparisons of triggering intervals.

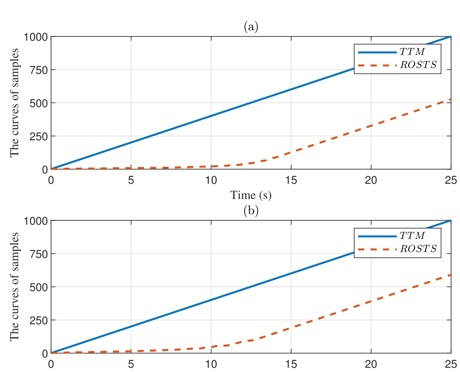

Fig. 6. Examples of different control strategies.

VI. Conclusion

This paper proposes an RL-based self-triggered control scheme combining GFHM adaptive fuzzy identification for optimal stabilization of unknown nonlinear NISs. The GFHM identifier leverages its universal approximation property to fit unknown nonlinear dynamics with low parameter overhead, laying a reliable foundation for optimal control design. The self-triggered control mechanism abandons continuous state monitoring, predicts update instants via software calculation, and reduces communication and computational consumption while avoiding Zeno behavior. By constructing a critic NN and tuning weights via RL gradient descent, the scheme approximates the HJB equation solution to obtain the optimal control law. Theoretical derivations confirm all closed-loop signals are stable in the sense of UUB, and numerical simulations validate the scheme's superior performance in balancing control effects and resource efficiency. Our follow-up work will extend this method to complex industrial systems subject to input saturation and external disturbances like [39], [40].

References

[1] Z. Pang, C. Xia, J. Zhang, Q. Han, and G. Liu, "A prediction-based approach realizing finite-time convergence of networked control systems," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 70, no. 7, pp. 2445–2449, Jul. 2023.

[2] L. Hallinan, J. D. Watson, and I. Lestas, "Inverse optimal control for passive network systems," *IEEE Trans. Autom. Control*, vol. 70, no. 10, pp. 6969–6976, Oct. 2025.

[3] A. Mirzaei and A. Ramezani, "Feasible cooperative linked-optimization-based distributed model predictive control for networked linear large-scale constrained systems: Application to power systems," *IEEE Trans. Control Netw. Syst.*, vol. 12, no. 1, pp. 776–786, Mar. 2025.

[4] O. MiguelEscrig and J. RomeroPérez, "Robustness study for networked event-based control system under regular quantization with hysteresis sampling," *IEEE Trans. Control Netw. Syst.*, vol. 12, no. 1, pp. 125–134, Mar. 2025.

[5] L. Qiu, B. He, C. Song, Z. Huang, J. Teng, and Z. Wu, "Event-triggered predictive control for Semi-Markov jumping networked control systems with random DoS attack," *IEEE Trans. Ind. Cyber- Phys. Syst.*, vol. 3, pp. 419–428, 2025.

[6] B. Li and X. Zhao, "Neural network-based adaptive sliding mode control for T-S fuzzy fractional order systems," *IEEE Trans. Circuits Syst. II: Exp. Briefs*, vol. 70, no. 12, pp. 4549–4553, Dec. 2023.

[7] Z. Yuan, C. Zhang, X. Shangguan, Y. Fan, and Y. He, "T–S fuzzy-based load frequency control of multiarea power system with hybrid delays: Performance analysis and improvement," *IEEE Trans. Fuzzy Syst.*, vol. 32, no. 12, pp. 6939–6950, Dec. 2024.

[8] J. Liu, W. Wang, B. Hu, J. Liu, X. Xie, and E. Tian, "Fuzzy-boosted event-triggered tracking control of unknown nonlinear networked systems: A PSO-driven RL approach," *IEEE Trans. Automat. Sci. Eng.*, vol. 22, pp. 7736–7747, 2025.

[9] B. Zhao, D. Liu, and C. Luo, "Reinforcement learning-based optimal stabilization for unknown nonlinear systems subject to inputs with uncertain constraints," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 31, no. 10, pp. 4330–4340, Oct. 2020.

[10] J. Liu, E. Ma, L. Zha, and E. Tian, "Distributed energy management for microgrids: When privacy preservation meets multiple mechanisms," *IEEE Trans. Consum. Electron.*, vol. 72, no. 1, pp. 1514–1523, Feb. 2026.

[11] F. Mahdavi Golmisheh and S. Shamaghdari, "Data-driven inverse reinforcement learning for heterogeneous optimal robust formation control," *IEEE Trans. Cybern.*, vol. 55, no. 5, pp. 2024–2037, May 2025.

[12] R. Lu et al., "Adaptive optimal surrounding control of multiple unmanned surface vessels via actor-critic reinforcement learning," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 36, no. 7, pp. 12173–12186, Jul. 2025.

[13] G. Wen and B. Li, "Optimized leader-follower consensus control using reinforcement learning for a class of second-order nonlinear multiagent systems," *IEEE Trans. Syst., Man, Cybern. Syst.*, vol. 52, no. 9, pp. 5546–5555, Sep. 2022.

[14] H. Zhang, X. Zhao, H. Wang, G. Zong, and N. Xu, "Hierarchical sliding-mode surface-based adaptive actor–critic optimal control for switched nonlinear systems with unknown perturbation," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 35, no. 2, pp. 1559–1571, Feb. 2024.

[15] J. Kim and I. Yang, "Maximum entropy optimal control of continuous-time dynamical systems," *IEEE Trans. Autom. Control*, vol. 68, no. 4, pp. 2018–2033, Apr. 2023.

[16] J. Zhao, Z. Wang, Y. Lv, C. Liu, and Z. Zhao, "Robust optimal prescribed performance control of adaptive cruise control systems with unknown dynamics," *IEEE Trans. Intell. Transp. Syst.*, vol. 26, no. 4, pp. 4757–4769, Apr. 2025.

[17] D. Yao, Y. Wang, H. Wei, and Y. Shi, "Quantized-observer-based event-triggered secure consensus for NMASs under DoS attacks," *IEEE Trans. Ind. Cyber- Phys. Syst.*, vol. 4, pp. 44–52, 2026.

[18] F. Mahdavi Golmisheh and S. Shamaghdari, "Optimal robust formation of multi-agent systems as adversarial graphical apprentice games with inverse reinforcement learning," *IEEE Trans. Automat. Sci. Eng.*, vol. 22, pp. 4867–4880, 2025.

[19] J. Liu, L. Jin, J. Liu, and E. Tian, "Encoding-decoding-enabled non-fragile secure control for nonlinear NCSs under cyber attacks," *IEEE Trans. Syst., Man, Cybern. Syst.*, early access Jun. 17, 2026.

[20] X. Yang and Q. Wei, "Adaptive dynamic programming for robust event-driven tracking control of nonlinear systems with asymmetric input constraints," *IEEE Trans. Cybern.*, vol. 54, no. 11, pp. 6333–6344, Nov. 2024.

[21] R. Goyal, M. Naveed Gul Mohamed, R. Wang, A. Sharma, and S. Chakravorty, "Information-state-based reinforcement learning for the control of partially observed nonlinear systems," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 36, no. 12, pp. 20386–20400, Dec. 2025.

[22] H. Dong, X. Zhao, and H. Yang, "Reinforcement learning-based approximate optimal control for attitude reorientation under state constraints," *IEEE Trans. Control Syst. Technol.*, vol. 29, no. 4, pp. 1664–1673, Jul. 2021.

[23] K. Zhang, S. Luo, H. Wu, and R. Su, "Data-driven tracking control for nonaffine yaw channel of helicopter via off-policy reinforcement learning," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 61, no. 3, pp. 7725–7737, Jun. 2025.

[24] X. Cheng, K. Liu, Y. Wang, and X. Sun, "Safety-critical event-triggered control for networked control systems under quantization and time-varying delay," *IEEE Trans. Control Netw. Syst.*, vol. 12, no. 3, pp. 2219–2230, Sep. 2025.

[25] K. Zhao, Y. Huang, and Y. Zhou, "Dynamic event-triggered adaptive asymptotic tracking control of uncertain systems with unified prescribed performance," *IEEE Trans. Automat. Sci. Eng.*, vol. 22, pp. 14752–14762, 2025.

[26] X. Xu, M. Hou, and G. Duan, "Dynamic periodic event-triggered adaptive prescribed performance control of uncertain semi-strict feedback systems with application," *IEEE Trans. Circuits Syst. I: Reg. Papers*, vol. 72, no. 8, pp. 4248–4260, Aug. 2025.

[27] D. Wang, Z. Yuan, A. Liu, Q. Hu, and J. Qiao, "Static and dynamic event-triggered control for disturbed interconnected systems through adaptive critic designs," *IEEE Trans. Syst., Man, Cybern. Syst.*, vol. 56, no. 2, pp. 878–892, Feb. 2026.

[28] B. Zhao, S. Zhang, and D. Liu, "Self-triggered approximate optimal neuro-control for nonlinear systems through adaptive dynamic programming," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 36, no. 3, pp. 4713–4723, Mar. 2025.

[29] D. Wang, L. Hu, L. Zhang, and J. Qiao, "Online optimal self-triggered sampling control with efficient adaptive critic learning," *IEEE Trans. Autom. Control*, vol. 70, no. 12, pp. 8438–8445, Dec. 2025.

[30] M. Wakaiki, "Self-triggered stabilization of contracting systems under quantization," *IEEE Trans. Autom. Control*, vol. 70, no. 8, pp. 5254–5269, Aug. 2025.

[31] C. Zhang et al., "Enhancing attitude tracking with self-learning control using Tanh-type learning intensity," *IEEE Trans. Automat. Sci. Eng.*, vol. 22, pp. 16976–16986, 2025.

[32] S. Xie and W. Sun, "Adaptive predefined-time safety learning control for switched multi-agent systems: An advanced encryption self-triggered algorithm," *IEEE Trans. Automat. Sci. Eng.*, vol. 22, pp. 2990–22002, 2025.

[33] L. Hu, D. Wang, J. Ren, and J. Qiao, "Online self-triggered transmission control with critic learning for discrete nonlinear systems," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 36, no. 10, pp. 18653–18665, Oct. 2025.

[34] H. Yin, W. Qi, J. H. Park, Z. Wu, and H. Yan, "Intelligent finite-time self-triggered control for fuzzy UMV systems with hybrid attacks," *IEEE Trans. Syst., Man, Cybern. Syst.*, vol. 55, no. 10, pp. 6558–6568, Oct. 2025.

[35] M. Z. Shahabadi, H. Atrianfar, G. B. Gharehpetian, and A. Yazdani, "Distributed self-triggered control of Islanded microgrids under switching communication topologies," *IEEE Trans. Ind. Informat.*, vol. 20, no. 12, pp. 13904–13915, Dec. 2024.

[36] K. Zhang, Q. Zhu, X. Xie, and R. Su, "A self-triggered adaptive dynamic programming scheme for unknown nonlinear systems by iterative model predictive process," *IEEE Trans. Automat. Sci. Eng.*, vol. 22, pp. 16121–16130, 2025.

[37] H. Zhang, H. Su, K. Zhang, and Y. Luo, "Event-triggered adaptive dynamic programming for non-zero-sum games of unknown nonlinear systems via generalized fuzzy hyperbolic models," *IEEE Trans. Fuzzy Syst.*, vol. 27, no. 11, pp. 2202–2214, Nov. 2019.

[38] H. Su, Y. Cui, H. Zhang, X. Xie, X. Liang, and J. Wang, "Adaptive secure finite-time optimal control of unknown nonlinear systems with state constraints via generalized fuzzy hyperbolic models," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 36, no. 9, pp. 15749–15763, Sep. 2025.

[39] J. Liu, Z. Liu, Y. Li, E. Tian, and C. Peng, "Privacy-protected formation control for unmanned surface vehicles: A neural predictor-based noncooperative game framework," *IEEE Trans. Veh. Technol.*, vol. 75, no. 7, pp. 12137–12147, Jul. 2026.

[40] E. Tian, Z. Wu, Y. Huang, and H. Chen, "A survey of cyber-security in cyber-physical systems: Attack strategies and protection mechanisms," *IEEE Trans. Ind. Cyber- Phys. Syst.*, vol. 4, pp. 392–404, 2026.