Reinforcement Learning-Boosted Event-Triggered Reliability Control for Uncertain CSTR System With Asymmetric Constraints

Jian Liu, *Member, IEEE*, Jiachen Ke, Jinliang Liu, *Member, IEEE*, Xiangpeng Xie, *Senior Member, IEEE*, and Engang Tian, *Member, IEEE*

***Abstract*—From the perspective of industrial production reliability, a robust event-triggered (ET) control strategy is presented for uncertain continuous stirred tank reactor (CSTR) system with asymmetric input constraints. To begin with, we propose a nonquadratic performance function to transform the robust control issue by constructing the relevant auxiliary dynamics. For effectively mitigating the pressure of data transmission and controller execution, a dynamic ET scheme (DETS) with an adjustable threshold function is adopted. Subsequently, we formulate the DETS-based Hamilton–Jacobi–Bellman (DET-HJB) equation according to optimality theory. In addition, a DETS-assisted reinforcement learning algorithm with a unique critic neural network can efficiently tackle the derived DET-HJB equation. Meanwhile, the corresponding critic weight is regulated on the basis of gradient descent technique and experience replay approach. By presenting a rigorous analysis under two situations, the uniform ultimate boundedness of auxiliary dynamics and weight approximation error can be ensured. Eventually, the feasibility of the proposed algorithm is demonstrated by experimental results of CSTR system.**

***Index Terms*—Asymmetric constraints, industrial production reliability, reinforcement learning (RL), robust event-triggered (ET) control, uncertain continuous stirred tank reactor (CSTR).**

I. Introduction

IN THE realm of chemical production, continuous stirred tank reactor (CSTR) system can be regarded as a paramount instrument, which has been extensively utilized in biopharmaceutical, petroleum production, and other industrial reactions. In the intricate chemical production process, the reaction temperature and exothermic rate in CSTR system are susceptible to the impacts of uncertainty and exogenous interference. Consequently, it will result in the aberrant reaction and product disqualification [1]. From the perspective of industrial production reliability [2], [3], [4], [5], [6], it is considerably vital to make reaction temperature unchanged. In the past few decades, several research results [7], [8], [9] have proposed the relative control scheme for CSTR system. For instance, Wang et al. [10] devised a decentralized fault tolerant control policy with application to CSTR system. In Yang and Wei's [11] work, a neurooptimal event-triggered (ET) control scheme was presented for CSTR dynamics in virtue of critic-only neural network (NN). Moreover, Zhou et al. [12] provided a data-based tracking controller for constrained CSTR dynamics by exploiting iterative adaptive dynamic programming (ADP) algorithm. On the ground of these discussions, the existing control approaches have achieved optimization of performance function. However, the uncertain characteristics of CSTR system with asymmetric control constraints have not been focused yet. In reality, asymmetric input constraints are required in industrial processes [12], [13] from the reliable perspective of temperature, voltages, and so on. To this end, we take efforts to investigate the robust control issue for uncertain CSTR system with asymmetric input constraints under reinforcement learning (RL) framework. It should be mentioned that the matched uncertainty is a special case of unmatched uncertainty in robust and reliable control issue [14]. Therefore, the unmatched situation will be explored in order to obtain more general consequences.

The aforesaid ADP [15], RL [16], and adaptive critic design [17] can be regarded as synonyms because they employ similar implementation spirits when confronting the optimal control issue. Initially, the traditional actor–critic (AC) dual NNs [18] were deployed to tackle this problem. The operating mechanism of the AC structure is composed of the following two aspects. 1) The actor NN delivers a strategy to the targeted system or environment. 2) The critic NN evaluates the performance/cost brought by this strategy and provides the feedback impacts to the actor NN. Afterwards, the critic-only NN [19], [20], [21] was put forward for nonlinear systems with measurable

Manuscript received 26 December 2023; revised 29 March 2024; accepted 27 May 2024. Date of publication 5 June 2024; date of current version 5 March 2025. This work was supported in part by the National Natural Science Foundation of China under Grant 62001210, Grant 62373252, Grant 61973152, and Grant 62373196, in part by the Natural Science Foundation of Jiangsu Province of China under Grant BK20230063, in part by the Postgraduate Research and Practice Innovation Program of Jiangsu Province under Grant KYCX23_1887, and in part by the Startup Foundation for Introducing Talent of NUIST under Grant 2024r063. Associate Editor: Z. Ye. *(Corresponding author: Jinliang Liu.)*

Jian Liu and Jiachen Ke are with the College of Information Engineering, Nanjing University of Finance and Economics, Nanjing 210023, China (e-mail: by.liujian@gmail.com; 2212602143@qq.com).

Jinliang Liu is with the School of Computer Science, Nanjing University of Information Science and Technology, Nanjing 210044, China (e-mail: liujinliang@vip.163.com).

Xiangpeng Xie is with the Institute of Advanced Technology, Nanjing University of Posts and Telecommunications, Nanjing 210023, China (e-mail: xiexiangpeng1953@163.com).

Engang Tian is with the School of Optical-Electrical and Computer Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China (e-mail: tianengang@163.com).

Digital Object Identifier 10.1109/TR.2024.3407090

state information. In contrast to the AC architecture, the critic-only NN not only reduces the potential approximation error originated from the actor NN, but also decreases the computational complexity during the learning process. Owing to these merits, the critic-only NN has been utilized to address various control problems, such as tracking control [12], [22], decentralized control [23], robust control [14], [24], and so on. In these existing results, the original control problem was converted into deriving the solution of Hamilton–Jacobi–Bellman (HJB) equation. To attain such a transformation, it is indispensable to construct an appropriate performance cost function involving the system state, the input features and the environmental factors. For example, Yang et al. [25] put forward a novel cost function with nonquadratic term by considering asymmetric input constraints. In Yang et al.'s [26] work, an estimated state-based performance function was designed for nonlinear systems subject to unknown interference. On account of the above-mentioned observations, the conventional performance cost function needs to be modified when confronted with different surroundings and production demand. Hence, it is of significance for devising a cost function to simultaneously cope with the unmatched uncertainty and asymmetric control constraints.

In industrial circumstance, the restricted computational and network resources constrain the high-frequency data exchange over the shared communication channel. In order to enhance the resource utilization, the burgeoning attentions have been witnessed in the research of ET mechanism [27], [28], [29], [30], [31]. Under the deployment of ET mechanism, the redundant data transmission can be efficiently decreased. To be concrete, the information exchanges will not occur if the artificially devised ET condition cannot be satisfied. In virtue of this characteristic, abundant optimal ET control algorithms have been proposed within RL framework. For example, Wang et al. [32] provided the optimal ET containment control strategy for multiagent systems on basis of AC structure. In Yang and He's [33] work, the ET H_∞ controller design was presented for constrained nonlinear systems by adopting a nonquadratic cost function. In addition, Wang et al. [34] investigated the RL algorithm-assisted ET control strategy for nonlinear interconnected systems. However, the aforementioned ET conditions are deduced with the aim of stabilizing the concerned systems rather than reducing the superfluous information transmission. Meanwhile, the so-called Zeno behavior [35] requires to be excluded via a strict mathematical proof in these ET schemes, which obviously increases the analysis complexity. Therefore, it is relatively imperative to deploy an effective ET scheme, which can lessen the communication pressure and obviate Zeno behavior without extra mathematical analysis.

Stimulated by the content previously, an RL-boosted robust ET control strategy is presented for uncertain CSTR system with asymmetric input constraints. In contrast to available results, we provide the following innovations of this article.

1) Focused on robustness and reliability, the results in [14] and [24] have successfully addressed the robust control issue for uncertain systems. Nevertheless, the control constraints were not considered such that the proposed strategies in [14] and [24] cannot be directly applied to the industrial production. Therefore, we construct a performance function by concurrently incorporating the features of asymmetric constraints and unmatched uncertainty.
2) A predesigned dynamic ET scheme (DETS) is applied to decrease the superfluous data transmission by utilizing a dynamically adjustable threshold function. In comparison with [23], [33], and [36], it is not required the Zeno behavior-related mathematical analysis attributed to the introduced minimum ET interval, which decreases analysis complexity and assure sampling rationality in practical engineering implementation.
3) Distinguished from [37] and [38], the equilibrium point of uncertain CSTR system is nonzero such that the proposed DETS-assisted RL algorithm can be applicable for more extensive dynamical systems. Furthermore, the AC structure adopted in [32] and [39] is simplified as critic-only NN to lessen the computational workload. To appropriately obtain the convergent critic weight, the historical state information is utilized rather than considering exploratory noise, which may cause the instability of CSTR system.

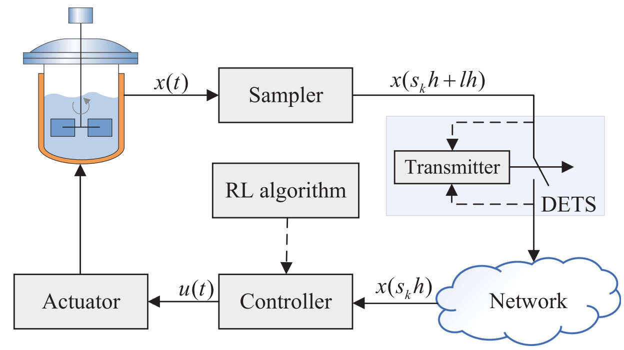


Fig. 1. Structure of uncertain CSTR system with reliable controller.

The rest of this article is organized as follows. Section II demonstrates the details of uncertain CSTR system, the transformation of robust control problem and the derivation of DETS-based HJB (DET-HJB) equation. In Section III, the DETS-assisted RL algorithm and its effectiveness analysis are presented. Moreover, the specific experimental results of uncertain CSTR system are given in Section IV. Finally, Section V concludes this article.

Notations: $g^\dagger(x)$ denotes the pseudoinverse of $g(x)$. Other adopted mathematical symbols in this article are standard.

II. Problem Statement and Transformation

In this section, the dynamics of uncertain CSTR system with asymmetric input constraints is presented in the first place. Afterward, the investigated robust control issue is converted into solving HJB equation via a modified cost function. Furthermore, the optimal ET control strategies under the employed DETS are deduced and the stability of uncertain CSTR system is guaranteed. To facilitate the next descriptions, Fig. 1 depicts the structure of uncertain CSTR system under reliable and robust control scheme.

A. Description of CSTR System

The dynamics of CSTR system [40] can be represented as

$$\begin{cases} \frac{dC_p}{dt_\tau} = \frac{C_I - C_p}{V} v_f - c_0 C_p e^{-\frac{E_a}{R_a T}} \\ \frac{dT}{dt_\tau} = \frac{T_f - T}{V} v_f - \Delta h \frac{c_0 C_p}{\rho_l C_l} e^{-\frac{E_a}{R_a T}} - \frac{T - T_i}{\rho_l V C_l} S_h A \end{cases} \tag{1}$$

where C_p is component concentration; C_I is input concentration; V is volume of reactor; v_f is feed velocity; c_0 is rate factor; E_a/R_a is activation energy term; T is reaction temperature; T_f is feed temperature; Δh is reaction heat; ρ_l is liquid density; C_l is heat capacity; T_i is inlet coolant temperature; and $S_h A$ is heat transfer term.

Define the auxiliary intermediate variables as follows:

$$\begin{cases} x_1 = \frac{C_I - C_p}{C_I}, \ x_2 = \frac{T - T_f}{T_f} \theta_a, \ \theta_a = \frac{E_a}{R_a T_f} \\ t = \frac{t_\tau v_f}{V}, \ u = \frac{T_i - T_f}{T_f} \theta_a, \ B_f = -\Delta h \frac{C_I \theta_a}{\rho_l C_l T_f} \\ D_f = \frac{c_0 V e^{-\theta_a}}{v_f}, \ \zeta = \frac{S_h A}{\rho_l v_f C_l}. \end{cases}$$

Then, the CSTR dynamics (1) is rewritten as

$$\frac{dx_1}{dt} = l_1(x_1, x_2), \ \frac{dx_2}{dt} = l_2(x_1, x_2) + \zeta u \tag{2}$$

where

$$l_1(x_1, x_2) = -x_1 + D_f(1 - x_1) e^{\frac{x_2 \theta_a}{\theta_a + x_2}}$$
$$l_2(x_1, x_2) = -(1 + \zeta) x_2 + B_f D_f (1 - x_1) e^{\frac{x_2 \theta_a}{\theta_a + x_2}}.$$

Let state vector $x(t) = [x_1^T, x_2^T]^T$, the functions $l(x(t)) = [l_1^T(x_1, x_2), l_2^T(x_1, x_2)]^T$, and $g(x(t)) = [0, \zeta^T]^T$. In line with the realistic environment, asymmetric constraints and unknown bounded uncertainty are taken into consideration. Then, the following uncertain CSTR system can be obtained:

$$\dot{x}(t) = l(x(t)) + g(x(t))u(t) + h(x(t))w(x(t)) \tag{3}$$

in which $x(t) \in \mathbb{R}^{n_x}$, $u(t) \in \mathcal{U} \subset \mathbb{R}^{n_u}$, and $w(x(t)) \in \mathbb{R}^{n_w}$ denote the measurable state, the control signal, and the unknown uncertainty, respectively. $\mathcal{U} = \{(u_1, u_2, \ldots, u_{n_u}) \in \mathbb{R}^{n_u} : \underline{u}_m \le u_\iota \le \overline{u}_m, \ \iota = 1, 2, \ldots, n_u\}$, from which $\underline{u}_m$ and $\overline{u}_m$ stand for the lower and upper bounds (LUBs) of u_ι with $|\underline{u}_m| \neq |\overline{u}_m|$. $l(x(t))$, $g(x(t))$, and $h(x(t))$ represent the given nonlinear functions with compatible dimensions. Meanwhile, the uncertainty is considered as the unmatched situation, i.e., $g(x(t)) \neq h(x(t))$.

For the bounded uncertainty $w(x(t))$ in (3), we assume that $||w(x(t))|| \le w_m(x(t))$ holds for a predetermined function $w_m(x(t))$ with $w(0) = 0$ and $w_m(0) = 0$. As demonstrated in Xue et al.'s [14] work, the uncertain term $h(x(t))w(x(t))$ is readily decomposed into the following form:

$$h(x(t))w(x(t)) = g(x(t))h_1(x(t)) + h_2(x(t))w(x(t)) \tag{4}$$

where

$$h_1(x(t)) = g^\dagger(x(t)) h(x(t)) w(x(t))$$
$$h_2(x(t)) = (I - g(x(t)) g^\dagger(x(t))) h(x(t)).$$

Remark 1: According to (3), it can be obtained that the equilibrium state of uncertain CSTR system is nonzero due to $l(0) \neq 0$ even though $u(t) = 0$ and $w(x(t)) = 0$. Apparently, it differs from the existing results in [37] and [38], which have supposed the system dynamics satisfying the condition $l(0) = 0$. Hence, the restrictive situation in [37] and [38] is relaxed in this article. On the other hand, the unmatched uncertainty and asymmetric input constraints are comprehensively considered such that the proposed control strategy can be applied to wide-ranging scenarios.

Assumption 1 ([14]): Supposing that $g(x(t))$, $h_1(x(t))$, and $h_2(x(t))$ are bounded by $||g(x(t))|| \le g_m$, $||h_1(x(t))|| \le h_{1m}$, and $||h_2(x(t))|| \le h_{2m}$, where $g_m > 0$, $h_{1m} > 0$, and $h_{2m} > 0$ stand for the given constants.

Definition 1 ([22]): If there exist a constant $\bar{\theta}$, a time instant $t_c = T(\bar{\theta}, x_0)$, and a compact set $\Omega \subset \mathbb{R}^{n_x}$ such that for $\forall t \ge t_0 + t_c$, $||x(t)|| \le \bar{\theta}$ holds with the initial state $x(t_0) = x_0 \in \Omega$, the nonlinear system is said to be uniformly ultimately bounded (UUB).

Under the aforementioned descriptions, this article endeavors to propose a robust control policy guaranteeing the UUB stability of uncertain CSTR system (3). However, it is worth mentioning that this main goal is relatively tough to attain owing to the asymmetric input constraints and the unmatched uncertainty. To obviate this difficulty, the design objective will be converted into obtaining the solution of optimal control issue over the corresponding auxiliary system.

B. Optimal Control Problem

In accordance with the dynamics (3), the following auxiliary system is constructed according to Zhang et al.'s [24] work:

$$\dot{x}(t) = l(x(t)) + g(x(t))u(t) + h_2(x(t))d(t) \tag{5}$$

in which $d(t)$ denotes a control variable in order to tackle the unmatched uncertainty. Thus, the control strategy pair of auxiliary system (5) can be described as $[u^T(t), \ d^T(t)]^T$.

Under the consideration of the unmatched uncertainty and asymmetric input constraints, we define the following cost function:

$$V(x(t)) = \int_t^\infty e^{-\sigma(s-t)} \left(\mathcal{Q}_1(x(s)) + \mathcal{Q}_2(x(s)) \right) ds \tag{6}$$

with

$$\mathcal{Q}_1(x(t)) = x^T(t) Q x(t) + \mathcal{R}(u(t)) + r_2 ||d(t)||^2$$
$$\mathcal{Q}_2(x(t)) = r_1 ||h_{1m}(x(t))||^2 + r_2 ||w_m(x(t))||^2$$

in which $\sigma > 0$ represents the discount factor and $e^{-\sigma(s-t)}$ can guarantee the convergence of cost function. $Q \in \mathbb{R}^{n_x \times n_x}$ denotes a predetermined positive definite matrix. $r_1 \ge 1$ and $r_2 > 0$ stand for the known parameters. Motivated by Yang and Wei's [25] work, the nonnegative function $\mathcal{R}(u(t))$ can be denoted by

$$\mathcal{R}(u(t)) = 2\alpha \sum_{i=1}^{n_u} \int_\beta^{u_i(t)} \Psi^{-1} \left(\frac{\tau_i - \beta}{\alpha} \right) d\tau_i \tag{7}$$

in which $\alpha = (\overline{u}_m - \underline{u}_m)/2$ and $\beta = (\overline{u}_m + \underline{u}_m)/2$. $\Psi^{-1}(\cdot)$ is the odd monotonic function with $\Psi^{-1}(0) = 0$. Facilitated by this characteristic, the hyperbolic tangent function is chosen as $\Psi(\cdot)$ in this article, i.e., $\Psi(\cdot) = \tanh(\cdot)$.

On basis of (6), the corresponding optimal cost function $V^*(x(t))$ is elicited as

$$V^*(x(t)) = \min_{u(t),d(t)\in\Phi(\Omega)} V(x(t)) \tag{8}$$

where $\Phi(\Omega)$ is the admissible set of all control strategies $u(t)$ and $d(t)$ over the compact set Ω.

By utilizing Bellman's optimality principle [33], $V^*(x(t))$ can be acquired through solving the following HJB equation:

$$\min_{u(t),d(t)\in\Phi(\Omega)} \mathcal{H}(\nabla V^*(x(t)), x(t), u(t), d(t)) = 0 \tag{9}$$

where $\nabla V^*(x(t))$ is the partial derivative of $V^*(x(t))$ with regard to $x(t)$. $\mathcal{H}(\nabla V^*(x(t)), x(t), u(t), d(t))$ represents Hamilton function with the form as

$$\begin{aligned}&\mathcal{H}(\nabla V^*(x(t)), x(t), u(t), d(t))\\&= (\nabla V^*(x(t)))^T(l(x(t)) + g(x(t))u(t) + h_2(x(t))d(t))\\&\quad + \mathcal{Q}_1(x(t)) + \mathcal{Q}_2(x(t)) - \sigma V^*(x(t)).\end{aligned} \tag{10}$$

Under the stationary condition [41], the following expressions can be obtained:

$$\begin{cases}\partial\mathcal{H}(\nabla V^*(x(t)), x(t), u(t), d(t))/\partial u(t) = 0\\ \partial\mathcal{H}(\nabla V^*(x(t)), x(t), u(t), d(t))/\partial d(t) = 0\end{cases} \tag{11}$$

and we can derive the following control strategies:

$$\begin{cases}u^*(x(t)) = -\alpha\Psi\left(\frac{1}{2\alpha}g^T(x(t))\nabla V^*(x(t))\right) + l_\beta\\ d^*(x(t)) = -\frac{1}{2r_2}h_2^T(x(t))\nabla V^*(x(t))\end{cases} \tag{12}$$

where $l_\beta = [\beta, \beta, \ldots, \beta]^T \in \mathbb{R}^{n_u}$ with β defined in the expression (7).

In light of the formulas (10) and (12), the HJB (9) is derived as

$$\begin{aligned}&(\nabla V^*(x(t)))^T(l(x(t))+g(x(t))u^*(x(t))+h_2(x(t))d^*(x(t)))\\&+ x^T(t)Qx(t) + \mathcal{R}(u^*(x(t))) + r_2||d^*(x(t))||^2\\&+ r_1||h_{1m}(x(t))||^2 + r_2||w_m(x(t))||^2 - \sigma V^*(x(t)) = 0.\end{aligned} \tag{13}$$

Apparently, the HJB (13) is obtained on basis of the time-driven control policies (12). Nevertheless, it is indubitable that such a mechanism will result in the heavy communication and computational workload. In order to overcome this problem, the DETS with an adjustable dynamic variable will be applied such that the redundant information transmission can be decreased. In the meantime, the HJB (13) will also be modified under the impact of DETS.

C. DETS-Based Optimal Control Policies

To enhance the network utilization, a DETS with a dynamic internal variable [42] is adopted to manage whether the data packets will be transmitted over the communication channel. In more detail, the sampled state information will be delivered and the control strategy $u(x(t))$ can be updated if the following inequality holds:

$$\eta(s_kh + lh)x^T(s_kh + lh)\Theta x(s_kh + lh) < e_k^T\Theta e_k \tag{14}$$

where $e_k = x(s_kh + lh) - x(s_kh)$ $(k = 0, 1, 2, \ldots;\ l = 1, 2, \ldots)$ denotes the gap between the current sampling time $s_kh + lh$ and the previous ET time s_kh. h represents the sampling period. $\Theta \in \mathbb{R}^{n_x\times n_x}$ is the positive definite matrix. In addition, the dynamic threshold function $\eta(s_kh + lh)$ is determined by

$$\eta(s_kh + lh) = \overline{\eta}_m + \frac{2}{\pi}(\underline{\eta}_m - \overline{\eta}_m)\mathrm{atan}(\epsilon e_k^T\Theta e_k) \tag{15}$$

where $\underline{\eta}_m$ and $\overline{\eta}_m$ stand for LUBs of $\eta(s_kh + lh)$, respectively. $\epsilon > 0$ is the known parameters to regulate the sensitivity of $\eta(s_kh + lh)$.

Based on the aforementioned content, the subsequent ET time $s_{k+1}h$ will be obtained via the following expression:

$$s_{k+1}h = s_kh + \min_{l\geq 1}\{lh|l \text{ satisfies the condition (14) }\}. \tag{16}$$

Under the utilization of the DETS and zero-order holder technique[11], the optimal control policies at each ET time can be represented as

$$\begin{cases}u^*(x(t)) = u^*(x(s_kh)),\ t \in [s_kh, s_{k+1}h)\\ d^*(x(t)) = -\frac{1}{2r_2}h_2^T(x(t))\nabla V^*(x(t))\end{cases} \tag{17}$$

where $u^*(x(s_kh)) = l_\beta - \alpha\Psi(\frac{1}{2\alpha}g^T(x(s_kh))\nabla V^*(x(s_kh)))$.

For brevity, x, x_k, and V_x^* stand for the notation of $x(t)$, $x(s_kh)$, and $\nabla V^*(x(t))$, respectively. Meanwhile, the DET-HJB equation is readily elicited as

$$\begin{aligned}&(V_x^*)^T(l(x) + g(x)u^*(x_k) + h_2(x)d^*(x)) + x^TQx\\&\quad + \mathcal{R}(u^*(x_k)) + r_2||d^*(x)||^2 + r_1||h_{1m}(x)||^2\\&\quad + r_2||w_m(x)||^2 - \sigma V^*(x) = 0,\ t \in [s_kh, s_{k+1}h).\end{aligned} \tag{18}$$

Remark 2: In general, the optimal control strategy $u^*(x)$ is computed at each ET time s_kh. For the control law $d^*(x)$, we take account of the scenario that $d^*(x)$ is updated at each sampling time according to Zhang et al.'s [24] work. Then, the optimal control strategy pair of auxiliary system (5) is described as $[u^{*T}(x_k),\ d^{*T}(x)]^T$. In addition, the DET-HJB (18) is also derived under this optimal control strategy pair. It is obviously witnessed that the DET-HJB (18) can be solved by fewer state information in comparison with the HJB (13). Thus, the communication and computational burden are alleviated under the applied DETS. For the convenient analysis, the interval $t \in [s_kh, s_{k+1}h)$ will not be emphasized in what follows unless noted otherwise.

D. Stability Analysis

Before proceeding, the following assumptions are presented to facilitate the theoretical analysis.

Assumption 2 ([43]): $||V^*(x)|| \leq V_m$ and $||V_x^*|| \leq V_{em}$, in which $V_m > 0$ and $V_{em} > 0$ denote the constants.

Assumption 3 ([44]): $u^*(x)$ is supposed to be Lipschitz continuous over the set $\Phi(\Omega)$ with the following condition:

$$||u^*(x) - u^*(x_k)|| \leq \mathcal{K}^*||x - x_k|| = \mathcal{K}^*||e_k|| \tag{19}$$

where $\mathcal{K}^*$ is a positive parameter.

Theorem 1: Under Assumptions 1–3, the given optimal cost function $V^*(x)$, and the optimal control policy $u^*(x_k)$ proposed in (17), the UUB stability of uncertain CSTR system (3) with asymmetric constraints can be ensured if the following inequality holds:

$$\lambda_{\min}(Q)\lambda_{\min}(\Theta) - \overline{\eta}_m(\mathcal{K}^*)^2\lambda_{\max}(\Theta) \ > \ 0. \tag{20}$$

Proof: According to the definition of cost function presented in (6), it can be concluded that $V^*(x) \geq 0$. Therefore, the Lyapunov function is constructed as $L_s(t) = V^*(x)$. Then, according to (4), its derivative is calculated as follows:

$$\begin{aligned}\dot{L}_s(t) &= (V_x^*)^T(l(x) + g(x)u^*(x_k) + h(x)w(x)) \\ &= (V_x^*)^T l(x) + (V_x^*)^T g(x)u^*(x_k) \\ &\quad + (V_x^*)^T g(x)h_1(x) + (V_x^*)^T h_2(x)w(x).\end{aligned} \tag{21}$$

In view of HJB (13), the following formula can be conveniently deduced:

$$\begin{aligned}(V_x^*)^T l(x) = &-(V_x^*)^T g(x)u^*(x) - (V_x^*)^T h_2(x)d^*(x) \\ &- x^T Qx - \mathcal{R}(u^*(x)) - r_2||d^*(x)||^2 \\ &- r_1||h_{1\,m}(x)||^2 - r_2||w_m(x)||^2 + \sigma V^*(x).\end{aligned} \tag{22}$$

On account of the expression (12), one has

$$\begin{cases}(V_x^*)^T g(x) = -2\alpha\Psi^{-T}\left((u^*(x) - l_\beta)/\alpha\right) \\ (V_x^*)^T h_2(x) = -2r_2(d^*(x))^T.\end{cases} \tag{23}$$

By adopting (22) and (23), the formula (21) is rewritten as

$$\begin{aligned}\dot{L}_s(t) = &\underbrace{2\alpha\Psi^{-T}\left((u^*(x) - l_\beta)/\alpha\right)(u^*(x) - u^*(x_k))}_{\pounds_1} \\ &\underbrace{-2\alpha\Psi^{-T}\left((u^*(x) - l_\beta)/\alpha\right)h_1(x)}_{\pounds_2} + r_2||d^*(x)||^2 \\ &- r_1||h_{1\,m}(x)||^2 - r_2||w_m(x)||^2 \underbrace{-2r_2(d^*(x))^T w(x)}_{\pounds_3} \\ &+ \sigma V^*(x) - x^T Qx - \mathcal{R}(u^*(x)).\end{aligned} \tag{24}$$

According to the inequality $2\tilde{a}^T\tilde{c} \leq ||\tilde{a}||^2 + ||\tilde{c}||^2$[23] and Assumption 3, we can get

$$\begin{aligned}\pounds_1 &\leq ||\alpha\Psi^{-1}((u^*(x) - l_\beta)/\alpha)||^2 + ||u^*(x) - u^*(x_k)||^2 \\ &\leq || - \frac{1}{2}g^T(x)V_x^*||^2 + (\mathcal{K}^*)^2||e_k||^2.\end{aligned} \tag{25}$$

Similarly, the following condition can also be derived:

$$\begin{aligned}\pounds_2 &\leq || - \alpha\Psi^{-1}((u^*(x) - l_\beta)/\alpha)||^2 + ||h_1(x)||^2 \\ &= ||\frac{1}{2}g^T(x)V_x^*||^2 + ||h_1(x)||^2.\end{aligned} \tag{26}$$

Likewise, noting that $||w(x(t))|| \leq w_m(x(t))$, it is simple to obtain that

$$\begin{aligned}\pounds_3 &\leq || - \sqrt{r_2}d^*(x)||^2 + ||\sqrt{r_2}w(x)||^2 \\ &\leq r_2||d^*(x)||^2 + r_2||w_m(x)||^2.\end{aligned} \tag{27}$$

By virtue of the formulas (24)–(27), the following condition is elicited:

$$\begin{aligned}\dot{L}_s(t) \leq &\ \sigma V^*(x) - x^T Qx - \mathcal{R}(u^*(x)) + \frac{1}{2}||g^T(x)V_x^*||^2 \\ &+ (\mathcal{K}^*)^2||e_k||^2 + ||h_1(x)||^2 - r_1||h_{1\,m}(x)||^2 \\ &+ 2r_2||d^*(x)||^2 \\ \leq &\ \sigma V^*(x) - \lambda_{\min}(Q)||x||^2 - \mathcal{R}(u^*(x)) \\ &+ \frac{1}{2}||g^T(x)V_x^*||^2 + (\mathcal{K}^*)^2||e_k||^2 + ||h_1(x)||^2 \\ &- r_1||h_{1\,m}(x)||^2 + \frac{1}{2r_2}||h_2^T(x)V_x^*||^2.\end{aligned} \tag{28}$$

In view of Assumptions 1 and 2, and noticing the facts that $\mathcal{R}(u^*(x)) \geq 0$ and $r_1 \geq 1$, the inequality (28) is deduced as

$$\begin{aligned}\dot{L}_s(t) \leq &\ \sigma V_m - \lambda_{\min}(Q)||x||^2 + (\mathcal{K}^*)^2||e_k||^2 \\ &+ \frac{1}{2}g_m^2 V_{em}^2 + \frac{1}{2r_2}h_{2\,m}^2 V_{em}^2.\end{aligned} \tag{29}$$

For $\forall t \in [s_k h, s_{k+1}h)$, the ET condition (14) is incapable to satisfy, which yields

$$\eta(s_k h + lh)x^T(s_k h + lh)\Theta x(s_k h + lh) \geq e_k^T\Theta e_k \tag{30}$$

and one can derive

$$||e_k||^2 \leq \overline{\eta}_m\frac{\lambda_{\max}(\Theta)}{\lambda_{\min}(\Theta)}||x(s_k h + lh)||^2. \tag{31}$$

On the basis of the aforementioned analysis, it is not difficult to obtain that

$$\begin{aligned}\dot{L}_s(t) \leq &\ \sigma V_m - \lambda_{\min}(Q)||x||^2 + (\mathcal{K}^*)^2\overline{\eta}_m\frac{\lambda_{\max}(\Theta)}{\lambda_{\min}(\Theta)}||x||^2 \\ &+ \frac{1}{2}g_m^2 V_{em}^2 + \frac{1}{2r_2}h_{2\,m}^2 V_{em}^2 \\ = &\ \sigma V_m + \frac{1}{2}g_m^2 V_{em}^2 + \frac{1}{2r_2}h_{2\,m}^2 V_{em}^2 \\ &- \left(\lambda_{\min}(Q) - \overline{\eta}_m(\mathcal{K}^*)^2\frac{\lambda_{\max}(\Theta)}{\lambda_{\min}(\Theta)}\right)||x||^2.\end{aligned} \tag{32}$$

Obviously, when the condition (20) satisfies, $\dot{L}_s(t) < 0$ holds if $x \notin \bar{\Omega}_x$ with $\bar{\Omega}_x$ denoted as

$$\bar{\Omega}_x = \left\{x:\ ||x|| \leq \sqrt{\frac{\Xi_1}{\Xi_2}}\right\} \tag{33}$$

where

$$\Xi_1 = \sigma V_m + \frac{1}{2}g_m^2 V_{em}^2 + \frac{1}{2r_2}h_{2\,m}^2 V_{em}^2$$

$$\Xi_2 = \lambda_{\min}(Q) - \overline{\eta}_m(\mathcal{K}^*)^2\frac{\lambda_{\max}(\Theta)}{\lambda_{\min}(\Theta)}.$$

Hence, the system state x converges to the ultimate bound being $\sqrt{\Xi_1/\Xi_2}$ while $t > t_0 + t_c$ with a time instant t_c. In line with Definition 1, the UUB stability of the uncertain CSTR system (3) with asymmetric constraints can be guaranteed under the optimal control policy $u^*(x_k)$ in (17). Therefore, this proof has been completed. ■

III. Implementation of RL Algorithm

Due to the existence of nonlinearity and asymmetric constraints, it is considerably difficult to obtain the analytical solution of DET-HJB (18). This section endeavors to acquire its approximate solution through an RL algorithm. Moreover, the effectiveness analysis of this algorithm will be conducted from two situations.

A. DETS-Assisted RL Algorithm

To lessen the computational pressure, only a single critic architecture is exploited to approximate the solution of DET-HJB (18). As stated in Yang and He's [33] work, the optimal performance function $V^*(x)$ in (8) is expressed as

$$V^*(x) = \mathcal{W}_c^T \varphi_c(x) + \varepsilon_c(x) \tag{34}$$

in which $\mathcal{W}_c \in \mathbb{R}^{n_c}$ represents the optimal weight. n_c is the quantity of neurons. $\varphi_c(x) = [\varphi_{c1}(x), \varphi_{c2}(x), \ldots, \varphi_{cn_c}(x)]^T$ denotes the activation function with mutually independent elements φ_{cs} $(s = 1, 2, \ldots, n_c)$. $\varepsilon_c(x)$ is the reconstruction error.

Conveniently, the partial derivative of (34) with regard to x is deduced as

$$V_x^* = \nabla\varphi_c^T(x)\mathcal{W}_c + \nabla\varepsilon_c(x). \tag{35}$$

Then, the optimal control policies in (17) are restated as

$$\begin{cases} u^*(x) = u^*(x_k) = -\alpha\Psi(X_1(x_k)) + \varepsilon_{u^*}(x_k) + l_\beta \\ d^*(x) = -\frac{1}{2r_2} h_2^T(x)\left(\nabla\varphi_c^T(x)\mathcal{W}_c + \nabla\varepsilon_c(x)\right) \end{cases} \tag{36}$$

where

$$X_1(x_k) = \frac{1}{2\alpha} g^T(x_k)\nabla\varphi_c^T(x_k)\mathcal{W}_c$$

$$\varepsilon_{u^*}(x_k) = -\frac{1}{2}\left(I_{n_u} - \mathcal{A}(p(x_k))\right) g^T(x_k)\nabla\varepsilon_c(x_k)$$

and $\mathcal{A}(p(x_k)) = \operatorname{diag}\{\Psi^2(p_q(x_k))\}$ $(q = 1, 2, \ldots, n_u)$ with $p(x_k) = [p_1(x_k), p_2(x_k), \ldots, p_{n_u}(x_k)]^T \in \mathbb{R}^{n_u}$ selected between $(1/(2\alpha))g^T(x_k)V_{x_k}^*$ and $X_1(x_k)$.

However, the optimal weight vector $\mathcal{W}_c$ is unavailable in realistic engineering such that the derived control strategy pair $[u^{*T}(x_k),\ d^{*T}(x)]^T$ in (36) will not be promptly implemented. In order to eliminate this obstacle, we employ an estimated weight $\hat{\mathcal{W}}_c$ to replace $\mathcal{W}_c$. Thereby, the corresponding approximation $\hat{V}(x)$ and its partial derivative $\hat{V}_x$ are obtained

$$\hat{V}(x) = \hat{\mathcal{W}}_c^T\varphi_c(x),\ \hat{V}_x = \nabla\varphi_c^T(x)\hat{\mathcal{W}}_c. \tag{37}$$

On basis of the expression (37), the estimated values of $u^*(x_k)$ and $d^*(x)$ are derived as follows:

$$\begin{cases} \hat{u}(x) = \hat{u}(x_k) = -\alpha\Psi(X_2(x_k)) + l_\beta \\ \hat{d}(x) = -\frac{1}{2r_2} h_2^T(x)\nabla\varphi_c^T(x)\hat{\mathcal{W}}_c \end{cases} \tag{38}$$

where

$$X_2(x_k) = \frac{1}{2\alpha} g^T(x_k)\nabla\varphi_c^T(x_k)\hat{\mathcal{W}}_c.$$

By combining the formulas (10), (37), and (38), the approximate Hamilton function $\hat{\mathcal{H}}(\hat{V}_x, x, \hat{u}(x_k), \hat{d}(x))$ is elicited as

$$\begin{aligned} &\hat{\mathcal{H}}(\hat{V}_x, x, \hat{u}(x_k), \hat{d}(x)) \\ &\quad = \hat{\mathcal{W}}_c^T\nabla\varphi_c(x)(l(x) + g(x)\hat{u}(x_k) + h_2(x)\hat{d}(x)) \\ &\qquad + x^TQx + \mathcal{R}(\hat{u}(x_k)) + r_2||\hat{d}(x)||^2 + r_1||h_{1\,m}(x)||^2 \\ &\qquad + r_2||w_m(x)||^2 - \sigma\hat{\mathcal{W}}_c^T\varphi_c(x). \end{aligned} \tag{39}$$

Defining the estimated error between $\hat{\mathcal{H}}(\hat{V}_x, x, \hat{u}(x_k), \hat{d}(x))$ and $\mathcal{H}(V_x^*, x, u^*(x_k), d^*(x))$ as e_c, it yields

$$\begin{aligned} e_c &= \hat{\mathcal{H}}(\hat{V}_x, x, \hat{u}(x_k), \hat{d}(x)) - \mathcal{H}(V_x^*, x, u^*(x_k), d^*(x)) \\ &= \hat{\mathcal{W}}_c^T\phi_c + \vartheta_c \end{aligned} \tag{40}$$

where

$$\phi_c = \nabla\varphi_c(x)(l(x) + g(x)\hat{u}(x_k) + h_2(x)\hat{d}(x)) - \sigma\varphi_c(x)$$

$$\begin{aligned} \vartheta_c &= x^TQx + \mathcal{R}(\hat{u}(x_k)) + r_2||\hat{d}(x)||^2 + r_1||h_{1\,m}(x)||^2 \\ &\quad + r_2||w_m(x)||^2. \end{aligned}$$

In order to let $e_c \to 0$, the proper tuning rule of $\hat{\mathcal{W}}_c$ in (40) is comparatively important. Thus, the gradient descent technique [24] can be utilized to minimize the target function. Besides, to efficiently make use of historical state information, the following modified target function is applied:

$$E_c = \underbrace{\frac{1}{2}e_c^Te_c}_{E_{c,1}} + \underbrace{\sum_{j=1}^{\mathcal{N}(t)}\frac{1}{2}e^{s_jh-t}e_{c,j}^Te_{c,j}}_{E_{c,2}} \tag{41}$$

in which $\mathcal{N}(t)$ stands for the number of ET times within the interval $[0, t]$. $e_{c,j}$ $(j = 1, 2, \ldots, \mathcal{N}(t))$ denotes the historical information at time s_jh with the following form:

$$e_{c,j} = \hat{\mathcal{W}}_c^T\phi_{c,j} + \vartheta_{c,j} \tag{42}$$

where

$$\begin{aligned} \phi_{c,j} &= \nabla\varphi_c(x_j)(l(x_j) + g(x_j)\hat{u}(x_k) + h_2(x_j)\hat{d}(x_j)) \\ &\quad - \sigma\varphi_c(x_j) \\ \vartheta_{c,j} &= x_j^TQx_j + \mathcal{R}(\hat{u}(x_k)) + r_2||\hat{d}(x_j)||^2 + r_1||h_{1\,m}(x_j)||^2 \\ &\quad + r_2||w_m(x_j)||^2. \end{aligned}$$

Then, the gradient descent technique is applied to E_c, and the tuning rule of $\hat{\mathcal{W}}_c$ can be obtained via

$$\begin{aligned} \dot{\hat{\mathcal{W}}}_c &= -\frac{\gamma_c}{(1+\phi_c^T\phi_c)^2}\frac{\partial E_{c,1}}{\partial\hat{\mathcal{W}}_c} - \sum_{j=1}^{\mathcal{N}(t)}\frac{\gamma_c}{(1+\phi_{c,j}^T\phi_{c,j})^2}\frac{\partial E_{c,2}}{\partial\hat{\mathcal{W}}_c} \\ &= -\frac{\gamma_c\phi_c}{(1+\phi_c^T\phi_c)^2}e_c - \sum_{j=1}^{\mathcal{N}(t)}\frac{\gamma_ce^{s_jh-t}\phi_{c,j}}{(1+\phi_{c,j}^T\phi_{c,j})^2}e_{c,j} \end{aligned} \tag{43}$$

in which $\gamma_c \in (0, 1)$ denotes the adjustable parameter involving the convergence of $\hat{\mathcal{W}}_c$.

Algorithm 1: DETS-Assisted RL Algorithm.

Initialization:
$\hat{u}(0),\ \hat{d}(0) \leftarrow$ Eq. (38) with $\hat{\mathcal{W}}(0)$ and $x(0)$, $t = t + h$;
Learning Stage:
while $t \leq T_{\max}$ **do**
  $x(t) \leftarrow$ Eq. (5) with $\hat{u}(t-h)$ and $\hat{d}(t-h)$;
  $\hat{\mathcal{W}}_c(t) \leftarrow$ Eq. (43) until $||\hat{\mathcal{W}}_c(t) - \hat{\mathcal{W}}_c(t-h)|| \leq \hat{\epsilon}_0$;
  $\hat{d}(t) \leftarrow$ Eq. (38) with $\hat{\mathcal{W}}_c(t)$;
  if *the ET condition (14) holds* **then**
    $\hat{u}(t) \leftarrow$ Eq. (38) with $\hat{\mathcal{W}}_c(t)$;
    Collecting $\phi_{c,j}, \vartheta_{c,j} \leftarrow$ Eq. (42);
  else
    $\hat{u}(t) = \hat{u}(t-h)$;
  end
  $t = t + h$;
end
Implementation Stage:
Resetting $t = 0$;
$\hat{u}(0) \leftarrow$ Eq. (38) with $\hat{\mathcal{W}}_c$ from learning stage, $t = t + h$;
while $t \leq T_{\max}$ **do**
  $x(t) \leftarrow$ Eq. (3) with $\hat{u}(t-h)$;
  if *the ET condition (14) holds* **then**
    $\hat{u}(t) \leftarrow$ Eq. (38) with $\hat{\mathcal{W}}_c$;
  else
    $\hat{u}(t) = \hat{u}(t-h)$;
  end
  $t = t + h$;
end

Denoting the weight estimation error $\tilde{\mathcal{W}}_c = \mathcal{W}_c - \hat{\mathcal{W}}_c$, it yields

$$\dot{\tilde{\mathcal{W}}}_c = -\gamma_c \left(\frac{\phi_c \phi_c^T}{(1+\phi_c^T \phi_c)^2} + \sum_{j=1}^{\mathcal{N}(t)} \frac{\phi_{c,j}\phi_{c,j}^T e^{s_j h - t}}{(1+\phi_{c,j}^T \phi_{c,j})^2} \right) \tilde{\mathcal{W}}_c + \frac{\gamma_c \phi_c}{(1+\phi_c^T\phi_c)^2}\epsilon_c + \sum_{j=1}^{\mathcal{N}(t)} \frac{\gamma_c \phi_{c,j} e^{s_j h - t}}{(1+\phi_{c,j}^T\phi_{c,j})^2}\epsilon_{c,j} \tag{44}$$

in which

$$\begin{aligned} \epsilon_c =& -\nabla \varepsilon_c^T(x)(l(x) + g(x)\hat{u}(x_k) + h_2(x)\hat{d}(x)) + \sigma\varepsilon_c(x) \\ \epsilon_{c,j} =& -\nabla \varepsilon_c^T(x_j)(l(x_j) + g(x_j)\hat{u}(x_k) + h_2(x_j)\hat{d}(x_j)) \\ &+ \sigma\varepsilon_c(x_j). \end{aligned}$$

To plainly demonstrate the whole execution process, the proposed DETS-assisted RL algorithm is displayed in Algorithm 1 in pseudocode form. From Algorithm 1, the approximately optimal control strategies $\hat{u}(x_k)$ and $\hat{d}(x)$ are acquired for auxiliary system (5) during the learning stage. Subsequently, $\hat{u}(x_k)$ with the convergent critic weight vector $\hat{\mathcal{W}}_c$ is utilized to tackle the robust control problem for uncertain CSTR system (3) during the implementation stage.

Remark 3: In the critic NN, the experience replay and gradient descent approach are utilized to obtain the tuning law of $\hat{\mathcal{W}}_c$ according to Yang and Wei's [25] work. Despite the similar structure stated in Yang and Wei's [25] work, the term $e^{s_j h - t}$ in the modified target function (41) places greater emphasis on the newer state information. In addition, considering the asymmetric control requirements, the deduced approximate optimal control policies (38) are more practical in comparison with Xue et al.'s [14] work.

Remark 4: The DETS-assisted RL algorithm includes initialization, learning stage and implementation stage with time window $T_{\max}$. For convenience of understanding, the relevant details of the proposed algorithm are elaborated from the following aspects.

1) For the selection of the initial weight $\hat{\mathcal{W}}_c(0)$, there exist two representative approaches for time-driven and event-assisted RL algorithms. As stated in Zhang et al.'s [24] work, the critic weight can be initialized to zero for time-driven RL algorithm under an essential assumption. For event-assisted RL algorithm, trial and error can be adopted to gain the initial admissible control strategy [25], [33]. Thus, the initial value $\hat{\mathcal{W}}_c(0)$ in this article is chosen via trial and error.
2) In the learning stage, the critic weight vector $\hat{\mathcal{W}}_c(t)$ can be trained online for auxiliary system (5) until the execution time exceeds $T_{\max}$ or satisfies the convergent condition $||\hat{\mathcal{W}}_c(t) - \hat{\mathcal{W}}_c(t-h)|| \leq \hat{\epsilon}_0$, where $\hat{\epsilon}_0 > 0$ represents the computational accuracy. Hence, the final critic weight vector $\hat{\mathcal{W}}_c$ is obtained.
3) By utilizing the critic weight $\hat{\mathcal{W}}_c$ originated from learning stage, the robust control strategy is acquired for uncertain CSTR system (3) with asymmetric input constraints. In virtue of DETS and zero-order holder technique, the ET control mechanism can be eventually obtained. It is noteworthy that the DETS is deployed at learning and implementation stages, which aim to economize the computational and network resources.

B. *Effectiveness Analysis of DETS-Assisted RL Algorithm*

This subsection dedicates to illustrate the validity of DETS-assisted RL algorithm. Therefore, we will prove that the weight estimation error and the auxiliary system (5) are UUB. In order to conveniently assist the analysis, the following assumption can be presented.

Assumption 4 ([45]): Supposing that positive parameters b_{ε^*}, $\nabla\varphi_{cm}$, $\nabla\varepsilon_{cm}$, and ϵ_{cm} exist such that $||\varepsilon_{u^*}(x_k)|| \leq b_{\varepsilon^*}$, $||\nabla\varphi_c(x)|| \leq \nabla\varphi_{cm}$, $||\nabla\varepsilon_c(x)|| \leq \nabla\varepsilon_{cm}$, $||\epsilon_c|| \leq \epsilon_{cm}$, and $||\epsilon_{c,j}|| \leq \epsilon_{cm}$.

Theorem 2: Under the approximate optimal control strategies $\hat{u}(x_k)$, $\hat{d}(x)$ in (38), the updating rule of $\hat{\mathcal{W}}_c$ in (43), and Assumptions 1–4, if the following conditions hold:

$$\begin{cases} \lambda_{\min}(Q)\lambda_{\min}(\Theta) - 2\overline{\eta}_m(\mathcal{K}^*)^2\lambda_{\max}(\Theta) > 0 \\ \frac{1}{2}\gamma_c\lambda_{\min}(\mathcal{L}) - \frac{1}{2r_2}h_{2m}^2\nabla\varphi_{cm}^2 > 0 \end{cases} \tag{45}$$

where $\mathcal{L} = \phi_c\phi_c^T/(1+\phi_c^T\phi_c)^2 + \sum_{j=1}^{\mathcal{N}(t)} \phi_{c,j}\phi_{c,j}^T e^{s_j h - t}/(1+\phi_{c,j}^T\phi_{c,j})^2$, then the auxiliary system (5) and the weight estimation error $\tilde{\mathcal{W}}_c$ are ensured to be UUB.

Proof: For auxiliary system (5), the following Lyapunov function is selected:

$$L(t) = L_1(t) + L_2(t) + L_c(t) \tag{46}$$

in which $L_1(t) = V^*(x)$, $L_2(t) = V^*(x_k)$, and $L_c(t) = \frac{1}{2}\tilde{\mathcal{W}}_c^T\tilde{\mathcal{W}}_c$. Owing to the utilization of DETS, the theoretical analysis is divided into the following two scenarios.

SI: For $t \in [s_k h, s_{k+1}h)$, the condition $\dot{L}_2(t) = 0$ can be obtained. On account of auxiliary dynamics (5), the derivative of $L_1(t)$ is calculated as follows:

$$\dot{L}_1(t) = (V_x^*)^T(l(x) + g(x)\hat{u}(x_k) + h_2(x)\hat{d}(x)). \tag{47}$$

By utilizing the expressions (22) and (23), $\dot{L}_1(t)$ in (47) can be derived as

$$\begin{aligned}\dot{L}_1(t) = \sigma V^*(x) + &\underbrace{2\alpha\Psi^{-T}\left((u^*(x) - l_\beta)/\alpha\right)\left(u^*(x) - \hat{u}(x_k)\right)}_{\varSigma_1}\\ &+ \underbrace{2r_2(d^*(x))^T(d^*(x) - \hat{d}(x))}_{\varSigma_2} - x^TQx - \mathcal{R}(u^*(x))\\ &- r_1||h_{1m}(x)||^2 - r_2||w_m(x)||^2 - r_2||d^*(x)||^2.\end{aligned} \tag{48}$$

Based on the inequality $2\tilde{a}^T\tilde{c} \le ||\tilde{a}||^2 + ||\tilde{c}||^2$ and the formula (23), the term $\varSigma_1$ in (48) is elicited as

$$\begin{aligned}\varSigma_1 &\le ||\alpha\Psi^{-1}((u^*(x) - l_\beta)/\alpha)||^2 + ||u^*(x) - \hat{u}(x_k)||^2\\ &= || - \frac{1}{2}g^T(x)V_x^*||^2 + ||u^*(x) - \hat{u}(x_k)||^2.\end{aligned} \tag{49}$$

Similarly

$$\begin{aligned}\varSigma_2 &\le ||\sqrt{r_2}d^*(x)||^2 + ||\sqrt{r_2}(d^*(x) - \hat{d}(x))||^2\\ &= r_2||d^*(x)||^2 + r_2||d^*(x) - \hat{d}(x)||^2.\end{aligned} \tag{50}$$

Applying the inequality $\frac{1}{2}||\tilde{a} + \tilde{c}||^2 \le ||\tilde{a}||^2 + ||\tilde{c}||^2$ [22], the conditions (36) and (38), the term $||u^*(x) - \hat{u}(x_k)||^2$ in (49) is tackled as

$$\begin{aligned}||u^*(x) - \hat{u}(x_k)||^2 &= ||u^*(x) - u^*(x_k) + u^*(x_k) - \hat{u}(x_k)||^2\\ &\le 2||u^*(x) - u^*(x_k)||^2 + 2||u^*(x_k) - \hat{u}(x_k)||^2\\ &\le 2||u^*(x) - u^*(x_k)||^2 + 4||\varepsilon_{u^*}(x_k)||^2\\ &\quad + 4||\alpha\Psi(X_2(x_k)) - \alpha\Psi(X_1(x_k))||^2\\ &\le 2||u^*(x) - u^*(x_k)||^2 + 4||\varepsilon_{u^*}(x_k)||^2\\ &\quad + 8\alpha^2\left(||\Psi(X_2(x_k))||^2 + ||\Psi(X_1(x_k))||^2\right).\end{aligned} \tag{51}$$

Denote $X_z(x_k) = [X_{z1}(x_k), X_{z2}(x_k), \ldots, X_{zn_u}(x_k)]^T \in \mathbb{R}^{n_u}$ $(z = 1, 2)$. In light of the truth that $|\tanh(\bar{y})| \le 1$ holds for $\forall\bar{y} \in \mathbb{R}$, it can be found that

$$||\Psi(X_z(x_k))||^2 = \sum_{i=1}^{n_u}\tanh^2(X_{zi}(x_k)) \le n_u. \tag{52}$$

In virtue of Assumptions 3, 4, and (52), the inequality (51) is converted to

$$||u^*(x) - \hat{u}(x_k)||^2 \le 2(\mathcal{K}^*)^2||e_k||^2 + 16\alpha^2 n_u + 4b_{\varepsilon^*}^2. \tag{53}$$

For the term $||d^*(x) - \hat{d}(x)||^2$ in (50), it is readily obtained from (36) and (38) that

$$\begin{aligned}&||d^*(x) - \hat{d}(x)||^2\\ &= || - \frac{1}{2r_2}h_2^T(x)\nabla\varphi_c^T(x)\tilde{\mathcal{W}}_c - \frac{1}{2r_2}h_2^T(x)\nabla\varepsilon_c(x)||^2\\ &\le 2||\frac{1}{2r_2}h_2^T(x)\nabla\varphi_c^T(x)\tilde{\mathcal{W}}_c||^2 + 2||\frac{1}{2r_2}h_2^T(x)\nabla\varepsilon_c(x)||^2\\ &= \frac{1}{2r_2^2}||h_2^T(x)\nabla\varphi_c^T(x)\tilde{\mathcal{W}}_c||^2 + \frac{1}{2r_2^2}||h_2^T(x)\nabla\varepsilon_c(x)||^2.\end{aligned} \tag{54}$$

Combining the formulas (48)–(50), (53), and (54), we can conclude

$$\begin{aligned}\dot{L}_1(t) \le\ &\sigma V^*(x) + || - \frac{1}{2}g^T(x)V_x^*||^2 + 2(\mathcal{K}^*)^2||e_k||^2\\ &+ 16\alpha^2 n_u + 4b_{\varepsilon^*}^2 + \frac{1}{2r_2}||h_2^T(x)\nabla\varphi_c^T(x)\tilde{\mathcal{W}}_c||^2\\ &+ \frac{1}{2r_2}||h_2^T(x)\nabla\varepsilon_c(x)||^2 - x^TQx - \mathcal{R}(u^*(x))\\ &- r_1||h_{1m}(x)||^2 - r_2||w_m(x)||^2.\end{aligned} \tag{55}$$

Under Assumptions 1, 2, and 4, noticing $\mathcal{R}(u^*(x)) \ge 0$, $r_1 \ge 1$ and $r_2 > 0$, it can be elicited that

$$\begin{aligned}\dot{L}_1(t) \le\ &\sigma V_m + \frac{1}{4}g_m^2 V_{em}^2 + 2(\mathcal{K}^*)^2||e_k||^2\\ &+ 16\alpha^2 n_u + 4b_{\varepsilon^*}^2 + \frac{1}{2r_2}h_{2m}^2\nabla\varphi_{cm}^2||\tilde{\mathcal{W}}_c||^2\\ &+ \frac{1}{2r_2}h_{2m}^2\nabla\varepsilon_{cm}^2 - \lambda_{\min}(Q)||x||^2.\end{aligned} \tag{56}$$

As for $\dot{L}_c(t)$, it can be deduced from (44) that

$$\begin{aligned}\dot{L}_c(t) &= \tilde{\mathcal{W}}_c^T\dot{\tilde{\mathcal{W}}}_c\\ &= -\gamma_c\tilde{\mathcal{W}}_c^T\left(\frac{\phi_c\phi_c^T}{(1 + \phi_c^T\phi_c)^2} + \sum_{j=1}^{\mathcal{N}(t)}\frac{\phi_{c,j}\phi_{c,j}^T e^{s_j h - t}}{(1 + \phi_{c,j}^T\phi_{c,j})^2}\right)\tilde{\mathcal{W}}_c\\ &\quad + \frac{\gamma_c\tilde{\mathcal{W}}_c^T\phi_c}{(1 + \phi_c^T\phi_c)^2}\epsilon_c + \sum_{j=1}^{\mathcal{N}(t)}\frac{\gamma_c\tilde{\mathcal{W}}_c^T\phi_{c,j}e^{s_j h - t}}{(1 + \phi_{c,j}^T\phi_{c,j})^2}\epsilon_{c,j}.\end{aligned} \tag{57}$$

According to the inequality $2\tilde{a}^T\tilde{c} \le \tilde{a}^T\tilde{a} + \tilde{c}^T\tilde{c}$, the condition (57) is written as

$$\begin{aligned}\dot{L}_c(t) \le &-\frac{\gamma_c}{2}\tilde{\mathcal{W}}_c^T\left(\frac{\phi_c\phi_c^T}{(1 + \phi_c^T\phi_c)^2} + \sum_{j=1}^{\mathcal{N}(t)}\frac{\phi_{c,j}\phi_{c,j}^T e^{s_j h - t}}{(1 + \phi_{c,j}^T\phi_{c,j})^2}\right)\tilde{\mathcal{W}}_c\\ &+ \frac{\gamma_c\epsilon_c^T\epsilon_c}{2(1 + \phi_c^T\phi_c)^2} + \sum_{j=1}^{\mathcal{N}(t)}\frac{\gamma_c\epsilon_{c,j}^T\epsilon_{c,j}e^{s_j h - t}}{2(1 + \phi_{c,j}^T\phi_{c,j})^2}.\end{aligned} \tag{58}$$

It should be noted that $s_j h - t \leq 0$, and then (58) under Assumption 4 is transformed as

$$\dot{L}_c(t) \leq -\frac{1}{2}\gamma_c \lambda_{\min}(\mathcal{L})||\tilde{\mathcal{W}}_c||^2 + \frac{1}{2}\gamma_c(1+\mathcal{N}(t))\epsilon_{cm}^2. \quad (59)$$

Substituting (31) into (56), and utilizing the formula (59), the following result can be derived:

$$\begin{aligned}\dot{L}(t) \leq{}& \sigma V_m + \frac{1}{4}g_m^2 V_{em}^2 + 2\overline{\eta}_m(\mathcal{K}^*)^2 \frac{\lambda_{\max}(\Theta)}{\lambda_{\min}(\Theta)}||x||^2 \\ &+ 16\alpha^2 n_u + 4b_{\varepsilon^*}^2 + \frac{1}{2r_2}h_{2m}^2 \nabla\varphi_{cm}^2 ||\tilde{\mathcal{W}}_c||^2 \\ &+ \frac{1}{2r_2}h_{2m}^2 \nabla\varepsilon_{cm}^2 - \lambda_{\min}(Q)||x||^2 \\ &- \frac{1}{2}\gamma_c\lambda_{\min}(\mathcal{L})||\tilde{\mathcal{W}}_c||^2 + \frac{1}{2}\gamma_c(1+\mathcal{N}(t))\epsilon_{cm}^2 \\ ={}& \Pi_1 - \Pi_2||x||^2 - \Pi_3||\tilde{\mathcal{W}}_c||^2 \end{aligned} \quad (60)$$

where

$$\begin{aligned}\Pi_1 ={}& \sigma V_m + \frac{1}{4}g_m^2 V_{em}^2 + 16\alpha^2 n_u + \frac{1}{2r_2}h_{2m}^2\nabla\varepsilon_{cm}^2 \\ &+ 4b_{\varepsilon^*}^2 + \frac{1}{2}\gamma_c(1+\mathcal{N}(t))\epsilon_{cm}^2 \\ \Pi_2 ={}& \lambda_{\min}(\Theta) - 2\overline{\eta}_m(\mathcal{K}^*)^2\frac{\lambda_{\max}(\Theta)}{\lambda_{\min}(\Theta)} \\ \Pi_3 ={}& \frac{1}{2}\gamma_c\lambda_{\min}(\mathcal{L}) - \frac{1}{2r_2}h_{2m}^2\nabla\varphi_{cm}^2.\end{aligned}$$

It is apparent to observe that $\dot{L}(t) < 0$ if the condition (45) satisfies, $x \notin \mathcal{C}_x$ or $\tilde{\mathcal{W}}_c \notin \mathcal{C}_{\tilde{\mathcal{W}}_c}$, in which $\mathcal{C}_x$ and $\mathcal{C}_{\tilde{\mathcal{W}}_c}$ are presented by

$$\begin{aligned}\mathcal{C}_x &= \left\{x: \ ||x|| \leq \sqrt{\frac{\Pi_1}{\Pi_2}}\right\} \\ \mathcal{C}_{\tilde{\mathcal{W}}_c} &= \left\{\tilde{\mathcal{W}}_c: \ ||\tilde{\mathcal{W}}_c|| \leq \sqrt{\frac{\Pi_1}{\Pi_3}}\right\}.\end{aligned} \quad (61)$$

Therefore, according to Definition 1 and Lyapunov theorem extension, the considered auxiliary system (5) and the estimated weight error $\tilde{\mathcal{W}}_c$ are UUB. In addition, their ultimate bounds are $\sqrt{\Pi_1/\Pi_2}$ and $\sqrt{\Pi_1/\Pi_3}$ defined in (61), respectively.

SII: The event is triggered, which yields $t = s_{k+1}h$. On account of the definition of $L(t)$ in (46), we can obtain its difference as

$$\begin{aligned}\Delta L(t) ={}& V^*(x(s_{k+1}h)) - V^*(x(s_{k+1}^- h)) + V^*(x(s_{k+1}h)) \\ &- V^*(x(s_k h)) + \frac{1}{2}\tilde{\mathcal{W}}_c^T(s_{k+1}h)\tilde{\mathcal{W}}_c(s_{k+1}h) \\ &- \frac{1}{2}\tilde{\mathcal{W}}_c^T(s_{k+1}^- h)\tilde{\mathcal{W}}_c(s_{k+1}^- h)\end{aligned} \quad (62)$$

where $x(s_{k+1}^- h) = \lim_{\varsigma\to 0^+} x(s_{k+1}h - \varsigma)$ with $\varsigma \in (0, s_{k+1}h - s_k h)$.

From the discussions in SI, $\dot{L}_1(t) + \dot{L}_c(t) < 0$ for $\forall t \in [s_k h, s_{k+1}h)$ can be concluded. This means that $L_1(t) + L_c(t)$ is monotonically decreasing over $[s_k h, s_{k+1}h)$. Meanwhile, it is evident that $L_1(t) + L_c(t)$ is continuous over $[s_k h, s_{k+1}h]$. Hence, we can obtain that $L_1(t) + L_c(t)$ is monotonically decreasing over $[s_k h, s_{k+1}h]$, which yields

$$L_1(s_{k+1}h) + L_c(s_{k+1}h) < L_1(s_{k+1}h - \varsigma) + L_c(s_{k+1}h - \varsigma). \quad (63)$$

Then, when $\varsigma \to 0^+$, one has

$$\begin{aligned}&L_1(s_{k+1}h) + L_c(s_{k+1}h) \\ &\quad < \lim_{\varsigma\to 0^+}(L_1(s_{k+1}h - \varsigma) + L_c(s_{k+1}h - \varsigma)) \\ &\quad = L_1(s_{k+1}^- h) + L_c(s_{k+1}^- h).\end{aligned} \quad (64)$$

According to the expression of $L_1(t)$ and $L_c(t)$ in (46), the inequality (64) can be restated as follows:

$$\begin{aligned}&V^*(x(s_{k+1}h)) + \frac{1}{2}\tilde{\mathcal{W}}_c^T(s_{k+1}h)\tilde{\mathcal{W}}_c(s_{k+1}h) \\ &- V^*(x(s_{k+1}^- h)) - \frac{1}{2}\tilde{\mathcal{W}}_c^T(s_{k+1}^- h)\tilde{\mathcal{W}}_c(s_{k+1}^- h) \ < \ 0.\end{aligned} \quad (65)$$

Since the UUB stability of auxiliary system (5) has been proved in SI, it is simple to indicate that $V^*(x(s_{k+1}h)) \leq V^*(x(s_k h))$. On basis of the aforementioned content, we get $\Delta L(t) < 0$ when $x \notin \mathcal{C}_x$ or $\tilde{\mathcal{W}}_c \notin \mathcal{C}_{\tilde{\mathcal{W}}_c}$. Taking the discussed scenarios into account, it can be attained that the auxiliary system (5) and the weight estimation error $\tilde{\mathcal{W}}_c$ are UUB. This proof has been finished. ■

IV. Experimental Results

To validate the feasibility of the developed DETS-based robust control policy, we will present the relevant simulation consequences in this position.

Motivated by Yang and Wei [11], the parameters of uncertain CSTR system (3) are set to be $B_f = 21.5$, $D_f = 0.036$, $\zeta = 25.2$, and $\theta_a = 28.5$. Let the initialization $x(0) = [1.5, -1.5]^T$, $u(t) \in \mathcal{U} = \{u \in \mathbb{R} : -0.4 \leq u \leq 0.7\}$ with $\underline{u}_m = -0.4$ and $\overline{u}_m = 0.7$. Considering the unmatched uncertainty inspired by Xue et al.'s [14] work, we select $h(x) = [1, -0.2]^T$ and $w(x) = d_1 x_1 \sin(d_2 x_2)$, where $d_1 \in [-2, 2]$, $d_2 \in [-2, 2]$. In light of the matrix theory and (4), the following conditions hold:

$$\begin{aligned}g^\dagger(x) &= (g^T(x)g(x))^{-1}g^T(x) = [0, 25.2] \\ ||h_1(x)|| &= ||g^\dagger(x)h(x)w(x)|| = 0 \triangleq h_{1m} \\ h_2(x) &= (I - g(x)g^\dagger(x))h(x) = [1, 0]^T \\ ||w(x)|| &\leq ||d_1 x_1|| \leq 2||x_1|| \triangleq w_m(x).\end{aligned}$$

Besides, the DETS-related parameters are selected as $\Theta = \mathrm{diag}\{0.4, 0.6\}$, $\underline{\eta}_m = 0.02$, $\overline{\eta}_m = 0.07$, $\epsilon = 30$, and $h = 0.01$s. For the parameters of cost function $V(x)$ defined in (6), we set $\sigma = 0.35$, $Q = \mathrm{diag}\{100, 100\}$, $r_1 = 1.5$, and $r_2 = 0.47$. As illustrated in Yang and He's [33] work, the non-negative function

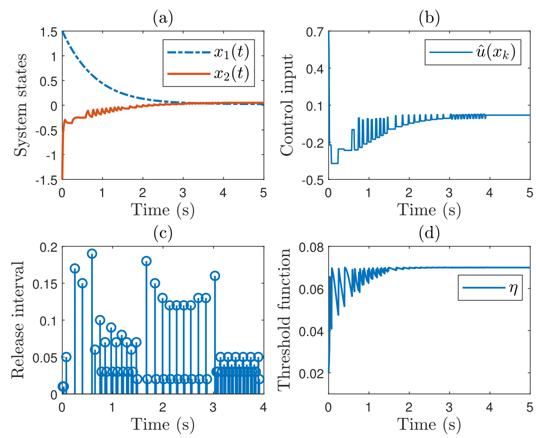


Fig. 2. Auxiliary system: (a) System states. (b) Approximate optimal control input. (c) DETS-based release interval. (d) Dynamic threshold function η.

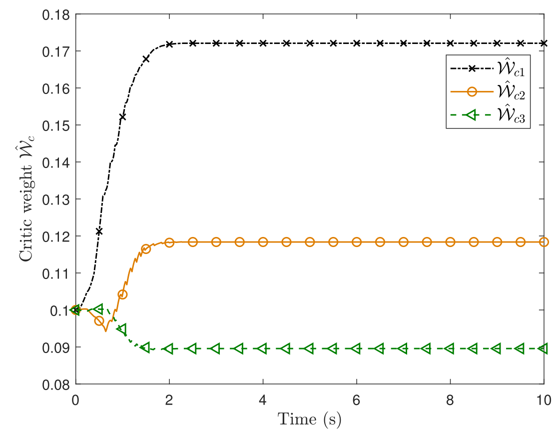


Fig. 3. Evolution of the critic weight $\hat{\mathcal{W}}_c$.

$\mathcal{R}(u(t))$ in (7) is represented as

$$\begin{aligned}\mathcal{R}(u(t)) &= 2\alpha \int_{\beta}^{u(t)} \tanh^{-1}((\lambda-\beta)/\alpha)d\lambda \\ &= 2\alpha(u(t)-\beta)\tanh^{-1}((u(t)-\beta)/\alpha) \\ &\quad + \alpha^2\ln(1-(u(t)-\beta)^2/\alpha^2).\end{aligned}$$

Meanwhile, it can be calculated from (7) that $\alpha = 0.55$ and $\beta = 0.15$. In the applied critic NN (37), we choose the learning rate $\gamma_c = 0.01$, the activation function $\varphi_c(x) = [x_1^2, x_1x_2, x_2^2]^T$, and the approximated weight $\hat{\mathcal{W}}_c = [\hat{\mathcal{W}}_{c1}, \hat{\mathcal{W}}_{c2}, \hat{\mathcal{W}}_{c3}]^T$. By adopting trial and error approach, we select the initial weight $\hat{\mathcal{W}}_c(0) = [0.1, 0.1, 0.1]^T$ to guarantee the initial control strategy being admissible.

Under the aforementioned parameters setting, the simulation process is conducted and the specific results are depicted in Figs. 2–7. By utilizing the proposed DETS-assisted RL algorithm, the critic weight vector $\hat{\mathcal{W}}_c$ is trained for auxiliary system (5) via the tuning law (43) and collecting the historical information data $\phi_{c,j}, \vartheta_{c,j}$, which are presented in Algorithm 1. During the learning process, the relevant trajectories of auxiliary system (5) are displayed in Fig. 2. Meanwhile, we choose the computational accuracy $\hat{\epsilon}_0 = 1 \times 10^{-5}$ and the corresponding evolution of critic weight $\hat{\mathcal{W}}_c$ is shown in Fig. 3. Upon the end of the learning stage, $\hat{\mathcal{W}}_c$ eventually converges to $[0.1721, 0.1184, 0.0896]^T$. From Fig. 2, it is plainly observed that the auxiliary system state and the approximate optimal control strategy can tend to the equilibrium point under the designed algorithm. In addition, the utilization of DETS also alleviate the computational pressure during the learning stage. Therefore, the effectiveness of Algorithm 1 is verified by the above-mentioned analysis.

Similar to Xue et al.'s [14] work, we randomly select the parameters $d_1 = 1.3658$ and $d_2 = 0.2516$. It can be witnessed from Fig. 4 that the states of uncertain CSTR system (3) are effectively stabilized by the designed optimal ET control scheme. Obviously, the states $x_1(t)$ and $x_2(t)$ tend to be nonzero condition. The phenomenon originates from the nonzero equilibrium point of uncertain CSTR system (3). Moreover, the approximate optimal ET control policy $\hat{u}(x_k)$ is plotted in Fig. 5, which reveals that the trajectory of $\hat{u}(x_k)$ does not exceed the asymmetric input bounds $\underline{u}_m$ and $\overline{u}_m$. It illustrates that the obstacle of asymmetric control constraints has been overcome for uncertain CSTR system (3). Meanwhile, $\hat{u}(x_k)$ also converges to nonzero value due to the existence of asymmetric constraints demonstrated in [46]. Besides, the DETS-based release instants and interval of uncertain CSTR system (3) are plotted in Fig. 6. The dynamic threshold function $\eta(s_kh + lh)$

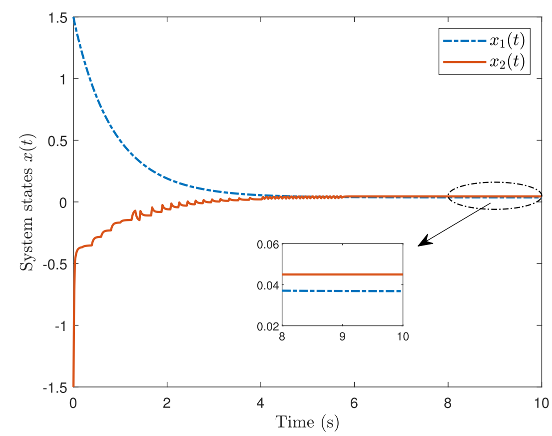


Fig. 4. States of uncertain CSTR system (3).

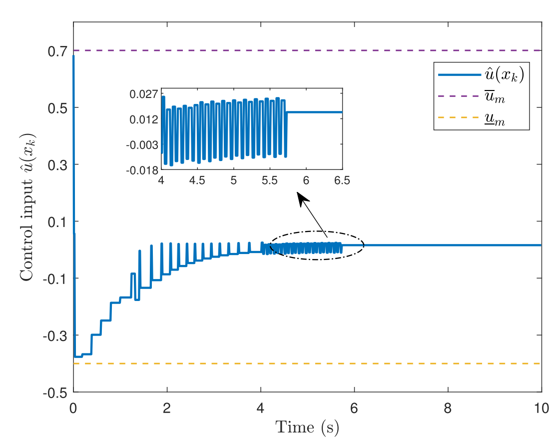


Fig. 5. Approximate optimal control $\hat{u}(x_k)$ of uncertain CSTR system (3).

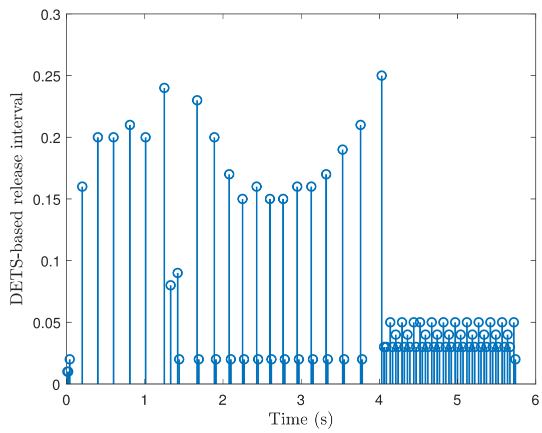


Fig. 6. DETS-based release interval of uncertain CSTR system (3).

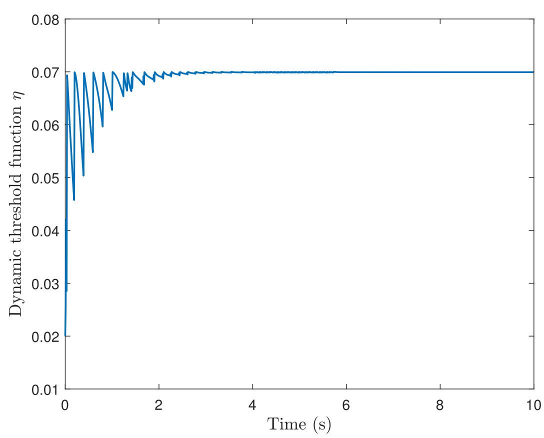


Fig. 7. Dynamic threshold function of uncertain CSTR system (3).

TABLE I
COMPARISON OF DIFFERENT TRIGGERED SCHEMES

Scheme	TDS	ETACL [25]	DETS
Triggered package	1000	114	83
Convergent time	4.98 s	5.39 s	5.74 s

is shown in Fig. 7. During the whole implementation stage, the function $\eta(s_k h + lh)$ is dynamically changing to regulate the ET threshold. It should be noticed that $\eta(s_k h + lh)$ eventually tends to be $\overline{\eta}_m$ when the uncertain CSTR system reaches the stable status about 5.74 s. Therefore, after 5.74 s, no events are triggered and no ET intervals can be witnessed in Fig. 6. On basis of the above-mentioned discussions, the proposed DETS-based approximate optimal control scheme is feasible and valid.

To demonstrate the advantages of DETS-assisted RL algorithm, we conduct the comparisons of triggered number under time-driven scheme (TDS) and ET adaptive critic learning (ETACL) approach proposed in Yang and Wei's [25] work. According to Table I, the better triggered performance is evidently obtained under DETS-assisted RL algorithm. To be specific, we actually adopt 83 sampled states to generate the approximate optimal control policy $\hat{u}(x_k)$. However, there are 1000 and 114 sampled states employed to calculate the control strategy under TDS and ETACL approach, respectively. On the other hand, the convergent difference of system states under these schemes is not apparent, which illustrates the applied DETS can balance the control and triggered performance. Hence, the practicability of the proposed control scheme is validated in virtue of the aforementioned comparisons.

V. CONCLUSION

This article presents an RL-boosted robust ET control policy for uncertain CSTR system with constrained input. In order to attain the industrial production reliability, asymmetric control constrains and unmatched uncertainty are simultaneously reflected in the designed performance function. Then, the investigated control problem can be converted into solving the targeted HJB equation. To lessen the computational burden, the predesigned DETS with a dynamic threshold function is adopted to appropriately regulate the data transmission. From the viewpoint of implementation, the optimal performance function is approximated by employing a DETS-assisted RL algorithm. Moreover, the auxiliary system states and the estimated weight error are proven to be UUB stability. In the end, the experimental results illustrate the practicability of the proposed robust control scheme. It is noteworthy that the applied gradient descent approach requires the appropriate initial weight vector. Thus, further research topics will involve the advanced weight tuning techniques within RL framework, such as adaptive moment estimation, particle swarm optimization and so on. In addition, how to develop an RL-boosted secure control scheme for large-scale networked systems against malicious attacks still deserves further investigations.

REFERENCES

[1] W. Zhou, J. Yi, L. Yao, and G. Chen, "Event-triggered optimal control for the continuous stirred tank reactor system," *IEEE Trans. Artif. Intell.*, vol. 3, no. 2, pp. 228–237, Apr. 2022.
[2] H. Wang and Z. S. Li, "Multi-area load frequency control in power system integrated with wind farms using fuzzy generalized predictive control method," *IEEE Trans. Rel.*, vol. 72, no. 2, pp. 737–747, Jun. 2023.
[3] Z. Ye, M. Revie, and L. Walls, "A load sharing system reliability model with managed component degradation," *IEEE Trans. Rel.*, vol. 63, no. 3, pp. 721–730, Sep. 2014.
[4] X. Cai, K. Shi, K. She, S. Zhong, Y. C. Soh, and Y. Yu, "Reliable sampling mechanism for Takagi-Sugeno fuzzy NCSs under deception cyberattacks for the application of the inverted pendulum system," *IEEE Trans. Rel.*, vol. 72, no. 3, pp. 1220–1231, Sep. 2023.
[5] P. Wen, Z. -S. Ye, Y. Li, S. Chen, P. Xie, and S. Zhao, "Physics-informed neural networks for prognostics and health management of lithium-ion batteries," *IEEE Trans. Intell. Veh.*, vol. 9, no. 1, pp. 2276–2289, Jan. 2024.
[6] R. Zhang and F. Gao, "A new synthetic minmax optimization design of H_∞ LQ tracking control for industrial processes under partial actuator failure," *IEEE Trans. Rel.*, vol. 69, no. 1, pp. 322–333, Mar. 2020.
[7] D. Li, D. Wang, L. Liu, and Y. Gao, "Adaptive finite-time tracking control for continuous stirred tank reactor with time-varying output constraint," *IEEE Trans. Syst., Man, Cybern. Syst.*, vol. 51, no. 9, pp. 5929–5934, Sep. 2021.
[8] Y. Wang, D. Li, and L. Liu, "IBLF-based adaptive finite-time control and modeling for continuous stirred tank reactor with output constraint," *J. Franklin Inst.*, vol. 359, no. 16, pp. 8669–8686, 2022.
[9] G. Wang, Q.-S. Jia, J. Qiao, J. Bi, and M. Zhou, "Deep learning-based model predictive control for continuous stirred-tank reactor system," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 8, pp. 3643–3652, Aug. 2021.

[10] Z. Wang, L. Liu, T. Li, and H. Zhang, "Minimum-learning-parameters-based adaptive neural fault tolerant control with its application to continuous stirred tank reactor," *IEEE Trans. Syst., Man, Cybern. Syst.*, vol. 50, no. 4, pp. 1275–1285, Apr. 2020.
[11] X. Yang and Q. Wei, "Adaptive critic designs for optimal event-driven control of a CSTR system," *IEEE Trans. Ind. Inform.*, vol. 17, no. 1, pp. 484–493, Jan. 2021.
[12] W. Zhou, H. Liu, H. He, J. Yi, and T. Li, "Neuro-optimal tracking control for continuous stirred tank reactor with input constraints," *IEEE Trans. Ind. Inform.*, vol. 15, no. 8, pp. 4516–4524, Aug. 2019.
[13] Y. Jiang, J. Fan, T. Chai, and F. L. Lewis, "Dual-rate operational optimal control for flotation industrial process with unknown operational model," *IEEE Trans. Ind. Electron.*, vol. 66, no. 6, pp. 4587–4599, Jun. 2019.
[14] S. Xue, B. Luo, and D. Liu, "Event-triggered adaptive dynamic programming for unmatched uncertain nonlinear continuous-time systems," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 7, pp. 2939–2951, Jul. 2021.
[15] H. Pan, C. Zhang, and W. Sun, "Fault-tolerant multiplayer tracking control for autonomous vehicle via model-free adaptive dynamic programming," *IEEE Trans. Rel.*, vol. 72, no. 4, pp. 1395–1406, Dec. 2023.
[16] J. Lu, Q. Wei, T. Zhou, Z. Wang, and F.-Y. Wang, "Event-triggered near-optimal control for unknown discrete-time nonlinear systems using parallel control," *IEEE Trans. Cybern.*, vol. 53, no. 3, pp. 1890–1904, Mar. 2023.
[17] Y. Huo, D. Wang, M. Li, and J. Qiao, "Decentralized event-triggered asymmetric constrained control through adaptive critic designs for nonlinear interconnected systems," *IEEE Trans. Syst., Man, Cybern. Syst.*, vol. 54, no. 1, pp. 391–402, Jan. 2024.
[18] X. Huo, H. R. Karimi, X. Zhao, B. Wang, and G. Zong, "Adaptive-critic design for decentralized event-triggered control of constrained nonlinear interconnected systems within an identifier-critic framework," *IEEE Trans. Cybern.*, vol. 52, no. 8, pp. 7478–7491, Aug. 2022.
[19] H. Zhang, Y. Liang, H. Su, and C. Liu, "Event-driven guaranteed cost control design for nonlinear systems with actuator faults via reinforcement learning algorithm," *IEEE Trans. Syst., Man, Cybern. Syst.*, vol. 50, no. 11, pp. 4135–4150, Nov. 2020.
[20] C. Zeng, B. Zhao, and D. Liu, "Fault tolerant control for a class of nonlinear systems with multiple faults using neuro-dynamic programming," *Neurocomputing*, vol. 553, 2023, Art. no. 126502.
[21] D. Wang and C. Mu, "Adaptive-critic-based robust trajectory tracking of uncertain dynamics and its application to a spring-mass-damper system," *IEEE Trans. Ind. Electron.*, vol. 65, no. 1, pp. 654–663, Jan. 2018.
[22] J. Liu, N. Zhang, Y. Li, X. Xie, E. Tian, and J. Cao, "Learning-based event-triggered tracking control for nonlinear networked control systems with unmatched disturbance," *IEEE Trans. Syst., Man, Cybern. Syst.*, vol. 53, no. 5, pp. 3230–3240, May 2023.
[23] X. Yang, Y. Zhu, N. Dong, and Q. Wei, "Decentralized event-driven constrained control using adaptive critic designs," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 33, no. 10, pp. 5830–5844, Oct. 2022.
[24] Q. Zhang, D. Zhao, and D. Wang, "Event-based robust control for uncertain nonlinear systems using adaptive dynamic programming," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 1, pp. 37–50, Jan. 2018.
[25] X. Yang and Q. Wei, "Adaptive critic learning for constrained optimal event-triggered control with discounted cost," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 1, pp. 91–104, Jan. 2021.
[26] Y. Yang et al., "Event-triggered output feedback control for a class of nonlinear systems via disturbance observer and adaptive dynamic programming," *IEEE Trans. Fuzzy Syst.*, vol. 31, no. 9, pp. 3148–3160, Sep. 2023.
[27] Y. Li, F. Song, J. Liu, X. Xie, and E. Tian, "Software defined event-triggering control for large-scale networked systems subject to stochastic cyber attacks," *IEEE Trans. Control Netw. Syst.*, vol. 10, no. 3, pp. 1531–1541, Sep. 2023.
[28] M. M. Hossain, C. Peng, H. -T. Sun, and S. Xie, "Bandwidth allocation-based distributed event-triggered LFC for smart grids under hybrid attacks," *IEEE Trans. Smart Grid*, vol. 13, no. 1, pp. 820–830, Jan. 2022.
[29] L. Zha, R. Liao, J. Liu, X. Xie, E. Tian, and J. Cao, "Dynamic event-triggered output feedback control for networked systems subject to multiple cyber attacks," *IEEE Trans. Cybern.*, vol. 52, no. 12, pp. 13800–13808, Dec. 2022.
[30] J. -W. Xing, C. Peng, Z. Cao, and W. -B. Xie, "Security-based control for networked unterval type-2 fuzzy systems with multiple cyber-attacks: An improved dynamic event-triggered scheme," *IEEE Trans. Fuzzy Syst.*, vol. 31, no. 8, pp. 2747–2760, Aug. 2023.
[31] E. Tian and C. Peng, "Memory-based event-triggering H_∞ load frequency control for power systems under deception attacks," *IEEE Trans. Cybern.*, vol. 50, no. 11, pp. 4610–4618, Nov. 2020.
[32] Z. Wang, X. Wang, and C. Zhao, "Event-triggered containment control for nonlinear multiagent systems via reinforcement learning," *IEEE Trans. Circuits Syst. II: Exp. Briefs*, vol. 70, no. 8, pp. 2904–2908, Aug. 2023.
[33] X. Yang and H. He, "Event-driven H_∞-constrained control using adaptive critic learning," *IEEE Trans. Cybern.*, vol. 51, no. 10, pp. 4860–4872, Oct. 2021.
[34] T. Wang, H. Wang, N. Xu, L. Zhang, and K. H. Alharbi, "Sliding-mode surface-based decentralized event-triggered control of partially unknown interconnected nonlinear systems via reinforcement learning," *Inf. Sci.*, vol. 641, 2023, Art. no. 119070.
[35] P. Tabuada, "Event-triggered real-time scheduling of stabilizing control tasks," *IEEE Trans. Autom. Control*, vol. 52, no. 9, pp. 1680–1685, Sep. 2007.
[36] S. Xue, B. Luo, D. Liu, and Y. Gao, "Event-triggered integral reinforcement learning for nonzero-sum games with asymmetric input saturation," *Neural Netw.*, vol. 152, pp. 212–223, 2022.
[37] C. Mu, K. Wang, and Z. Ni, "Adaptive learning and sampled-control for nonlinear game systems using dynamic event-triggering strategy," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 33, no. 9, pp. 4437–4450, Sep. 2022.
[38] Q. Wei, D. Liu, Q. Lin, and R. Song, "Adaptive dynamic programming for discrete-time zero-sum games," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 4, pp. 957–969, Apr. 2018.
[39] D. Wang, W. Fan, A. Liu, and J. Qiao, "Decentralized optimal neurocontroller design for mismatched interconnected systems via integral policy iteration," *IEEE Trans. Circuits Syst. II: Exp. Briefs*, vol. 71, no. 2, pp. 687–691, Feb. 2024.
[40] A. Uppal, W. Ray, and A. Poore, "On the dynamic behavior of continuous stirred tank reactors," *Chem. Eng. Sci.*, vol. 29, no. 4, pp. 967–985, 1974.
[41] W. Rudin, *Principles of Mathematical Analysis*, vol. 3. New York, NY, USA: McGraw-Hill, 1976.
[42] J. Liu, N. Zhang, L. Zha, X. Xie, and E. Tian, "Reinforcement learning-based decentralized control for networked interconnected systems with communication and control constraints," *IEEE Trans. Automat. Sci. Eng.*, 2023, to be published, doi: 10.1109/TASE.2023.3300917.
[43] M. Abu-Khalaf, J. Huang, and F. L. Lewis, *Nonlinear* H_2/H_∞ *Constrained Feedback Control: A Practical Design Approach Using Neural Networks*. Berlin, Germany: Springer, 2006.
[44] B. Zhao and D. Liu, "Event-triggered decentralized tracking control of modular reconfigurable robots through adaptive dynamic programming," *IEEE Trans. Ind. Electron.*, vol. 67, no. 4, pp. 3054–3064, Apr. 2020.
[45] D. Yang, T. Li, H. Zhang, and X. Xie, "Event-trigger-based robust control for nonlinear constrained-input systems using reinforcement learning method," *Neurocomputing*, vol. 340, pp. 158–170, 2019.
[46] X. Yang and B. Zhao, "Optimal neuro-control strategy for nonlinear systems with asymmetric input constraints," *IEEE/CAA J. Automatica Sinica*, vol. 7, no. 2, pp. 575–583, Mar. 2020.

Jian Liu (Member, IEEE) received the Ph.D. degree in signal and information processing from the College of Communication and Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing, China, in 2018.

He is currently an Associate Professor with the College of Information Engineering, Nanjing University of Finance and Economics, Nanjing, China. He has authored or coauthored more than 40 papers in refereed international journals. His research interests include cyberphysical systems, networked control systems, complex dynamical networks, intelligent optimization algorithms, and network security.

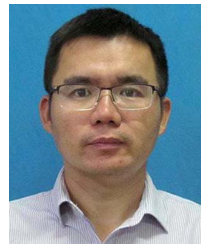
Dr. Liu was a recipient of the reward of Outstanding Reviewer for many international journals, such as *Computer Networks* and *Journal of Computational Science*. Since Aug. 2022, he has been a Guest Editor for the special issue Fuzzy Modeling and Fuzzy Control Systems in Mathematics.

Jiachen Ke received the B.Ed. degree in science education from the School of Science, Huzhou University, Huzhou, China, in 2022. He is currently working toward the M.S. degree in computer science and technology with the College of Information Engineering, Nanjing University of Finance and Economics, Nanjing, China.

His research interests include fuzzy control, learning-based optimal control methods, and networked control systems.

Jinliang Liu (Member, IEEE) received the Ph.D. degree in control theory and control engineering from the School of Information Science and Technology, Donghua University, Shanghai, China, in 2011.

He is currently a Professor with the School of Computer Science, Nanjing University of Information Science and Technology, Jiangsu, China. From 2013 to 2016, he was a Postdoctoral Research Associate with the School of Automation, Southeast University, Nanjing, China. From 2016 to 2017, he was a Visiting Researcher/Scholar with the Department of Mechanical Engineering, University of Hong Kong, Hong Kong. From 2017 to 2018, he was a Visiting Scholar with the Department of Electrical Engineering, Yeungnam University, Gyeongsan, South Korea. His research interests include networked control systems, complex dynamical networks, and time-delay systems.



Xiangpeng Xie (Senior Member, IEEE) received the B.S. and Ph.D. degrees in engineering from Northeastern University, Shenyang, China, in 2004 and 2010, respectively.

He is currently a Professor with the Institute of Advanced Technology, Nanjing University of Posts and Telecommunications, Nanjing, China. From 2010 to 2014, he was a Senior Engineer with the Metallurgical Corporation of China Ltd., Beijing, China.His research interests include fuzzy modeling and control synthesis, state estimations, optimization in process industries, and intelligent optimization algorithms.

Prof. Xie is an Associate Editor for IEEE TRANSACTIONS ON FUZZY SYSTEMS, IEEE TRANSACTIONS ON CYBERNETICS, IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS, *International Journal of Fuzzy Systems*, and *International Journal of Control, Automation, and Systems*.

Engang Tian (Member, IEEE) received the B.S. degree in mathematics from Shandong Normal University, Jinan, China, in 2002, the M.Sc. degree in operations research and cybernetics from Nanjing Normal University, Nanjing, China, in 2005, and the Ph.D. degree in control theory and control engineering from Donghua University, Shanghai, China, in 2008.

From 2011 to 2012, he was a Postdoctoral Research Fellow with the Hong Kong Polytechnic University, Hong Kong. From 2015 to 2016, he was a Visiting Scholar with the Department of Information Systems and Computing, Brunel University London, Uxbridge, U.K. From 2008 to 2018, he was an Associate Professor and then a Professor with the School of Electrical and Automation Engineering, Nanjing Normal University. In 2018, he was appointed as an Eastern Scholar by the Municipal Commission of Education, Shanghai, and joined the University of Shanghai for Science and Technology, Shanghai, China, where he is currently a Professor with the School of Optical-Electrical and Computer Engineering. He has authored or coauthored more than 100 papers in refereed international journals. His research interests include networked control systems, cyberattack, as well as nonlinear stochastic control and filtering.