# Reinforcement Learning-Based Event-Triggered Optimal Control of Power Systems With Control Input Saturation

Zhou Gu, *Senior Member, IEEE*, Ruiyan Cao, and Engang Tian, *Senior Member, IEEE*

***Abstract*—In this article, a reinforcement learning (RL)-based event-triggered guaranteed cost control (GCC) is developed for power systems with control input saturation (CIS). An event-triggered mechanism with a balance factor is developed to optimize both control performance of power systems and computation efficiency. To obtain an online solution to the corresponding modified Hamilton–Jacobi–Bellman equation for the power system, a critic neural network is employed. In contrast to RL-based methods employing dual-network setups, the single neural network not only conserves computational resources but also eliminates the need for an initial admissible control. For accelerating convergence toward optimal solutions, a new adaptive weight tuning law is constructed. Control performance of power systems with CIS and limited computation power is ensured by the proposed RL-based optimal GCC strategy. Simulation results validate the effectiveness of the proposed methodology.**

***Index Terms*—Event-triggered mechanism, power systems, reinforcement learning (RL).**

## I. Introduction

CURRENTLY, the power system is advancing toward intelligence, with renewable energy assuming a progressively vital role. However, its inherent instability and intermittent nature pose significant challenges. The integration of renewable energy on a large scale may lead to grid frequency fluctuations and voltage instability, thereby increasing operational risks in the power system [1]. Consequently, precise control of the synchronous generator system (SGS) becomes imperative to maintain the system's stability [2]. This control mechanism not only mitigates the potential for power outages and disruptions but also guarantees a dependable power supply within the system, ensuring users receive uninterrupted and stable power services.

The inherent complexity of SGSs systems, characterized by their nonlinear and dynamic attributes [3], poses challenges for conventional control techniques. Reinforcement learning (RL) addresses this challenge by leveraging system states and performance, enabling dynamic calibration and optimization of control strategies in response to system fluctuations and external perturbations. In contrast to conventional methods, RL demonstrates superior adaptability to diverse operational conditions, thereby facilitating precise control and enhancing the reliability and stability of SGSs. Recently, numerous RL-based [4], [5], [6] and ADP-based algorithms [7], [8], [9] have been devised to tackle the optimal control problem. For instance, in [8], the hybrid iteration method that combines both policy iteration and value iteration is developed. This approach mitigates the need for an admissible control policy (ACP) and achieves more rapid convergence to the optimal solution compared to sole reliance on value iteration. Meanwhile, the incorporation of both robustness and optimality within nonlinear control has attracted significant attention from numerous researchers [10], [11]. In [12], robust off-policy RL-based method was adopted to address nonlinear two-player Stackelberg game problem amid the existence of external perturbations. A novel RL-based approach was investigated in [13] to ensure the stability of dynamic systems facing unmatched uncertainties. This is achieved through the construction of an auxiliary system, facilitating the derivation of an optimal control law simultaneously.

Recently, event-triggering mechanism (ETM) has been widely used in the control field, which aims to reduce communication frequencies while maintaining adequate control performance, thereby lowering hardware requirements and conserving computing resources within control systems [14]. The releasing packets are selected from the sampling packets balancing the control and network bandwidth/computation power requirement by constructing a novel ETM in [15] and[16]. In [17] and[18], the authors extended the results of existing static ETMs through dynamically updating the threshold, enhancing triggering efficiency. Nevertheless, formulating an efficient ETM within

Received 26 April 2024; revised 8 August 2024; accepted 5 October 2024. This work was supported in part by the National Natural Science Foundation of China under Grant 62273183 and Grant 62103193; in part by the Natural Science Foundation of Jiangsu Province of China under Grant BK20231288; and in part by the Startup Funding of Anhui Polytechnic University under Grant 2024YQQ001. Paper no. TII-24-1798. *(Corresponding author: Zhou Gu.)*

Zhou Gu is with the College of Mechanical and Electronic Engineering,Nanjing Forestry University, Nanjing 210037, China, and also with the School of Electrical Engineering,Anhui Polytechnic University, Wuhu 241000, China (e-mail: gzh1808@163.com).

Ruiyan Cao is with the College of Mechanical and Electronic Engineering,Nanjing Forestry University, Nanjing 210037, China (e-mail: cry2514509931@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).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TII.2024.3485724.

Digital Object Identifier 10.1109/TII.2024.3485724

the RL framework poses significant challenges. While some scholars have initiated a preliminary exploration of RL-based event-triggered mechanisms, for example, an RL-based method in [19] employing ETM was devised to address challenge of zero-sum game in systems with partial unknowns, by which it reduces computational complexity while ensuring bounded states and neural network weight errors. Currently, numerous ETM designs based on RL utilize threshold values formulated as absolute error, such as in [20] and [21]. It is demonstrated that ETMs employing absolute error threshold exhibit limited adaptability. Moreover, it is hard to find a suitable threshold for the system with distinguished different states. Therefore, exploring alternative ETMs for nonlinear systems becomes necessary, leveraging insights from existing results for linear systems [22]. To our understanding, research in this area under the RL framework remains limited, thus motivating the current investigation in this study.

The control input of the SGS often faces limitations stemming from electromagnetic or mechanical constraints [23]. Traditional control methodologies, especially those based on the linear systems theory, often struggle to address these challenges effectively. Control input saturation (CIS) is among these issues. Failure to adequately address CIS within SGSs can readily lead to degradation of system performance, and in severe cases, instability [24]. RL, known for its robust optimization and control capabilities, offers a solution for effectively handling the CIS problem. In [25], asymmetric actuator saturation was investigated by employing novel policy iteration within the structure of RL. Qin et al. [26] investigated an RL-based control strategy for systems with CIS and collision avoidance constraints. An optimal tracking control scheme that combines admittance adaptation and method was developed in [27] for robot systems encountering environmental interaction and actuator saturation. Therefore, under the RL-based method, the control input is optimized while prioritizing system stability, ultimately maximizing the performance of the SGS, thereby enhancing system reliability, which stands as another crucial motivation driving our research endeavor.

Building on the aforementioned analysis, this article centers on event-triggered guaranteed cost control (GCC) for SGSs with CISs under the RL framework. The contributions of this study can be outlined as follows

1) The relative ETM that unifies the ETMs for linear and nonlinear systems is developed to handle the huge load of the computation. In contrast to the traditional ETM in absolute form, the proposed relative ETM is more reasonable for the practical application. Furthermore, the proposed ETM is incorporated with the balance factor to balance the control and computation efficiency.
2) Compared with the RL-based control method in [28], which utilizes a dual-network setup, our approach employs a single neural network to conserve computational resources. In addition, an extra term is integrated into the weight tuning law to reduce the necessity of ACP and accelerate convergence toward the optimal solution.
3) Different from most of the existing works, such as in [29], the optimal control policy in this study considers both the limited computation power and the CIS for SGSs.

The rest of this article are organized as follows. Section II presents the system description. Section III presents the transformation from robust GCC problem into optimal control problem for SGSs and introduces the event-based control design for the SGSs. In Section IV, the implementation of neural network, the main results of the RL-based algorithm and stability analysis under ETM are presented. The simulation conducted in Section V verifies the effectiveness of the proposed method. Finally, Section VI concludes this article.

Notations: $\mathbb{R}^m$ denotes Euclidean space comprised of $m$-dimensional vectors. $\mathbb{R}^{m\times n}$ is space of matrix is $m\times n$-dimensional. $T$ represents matrix transposition, which is transpose. $\lambda_{\min}$ and $\lambda_{\max}$ state the least and largest eigenvalue of matrix, respectively. $\max(\cdot)$ means the largest one in all elements. The operation $\{x|Q\}$ means $\{x|Q\}=x^TQx$. Let $\mathscr{A}(\mathfrak{Q})$ represent the set composed of all ACP, where $\mathfrak{Q}$ is a compact set belongs to $\mathbb{R}^m$. $\|\cdot\|$ refers to the 2-norm of vector or the matrix. $\nabla(\cdot)$ represents the gradient operator with respect to the variable $x\in\mathbb{R}^m$.

## II. System Description

Consider an SGS with the following model [3], [30]:

$$\begin{cases}\dot{\delta}=\omega\\ \dot{\omega}=-\frac{D}{2H}\omega+\mathcal{H}(T_G\mu-T_GK_G\omega-P_e)\\ \dot{E}'_q=-\sigma\frac{x_s}{x'_s}E'_q+\sigma p\cos(\delta)+\sigma E_f\end{cases}\tag{1}$$

where $\mathcal{H}=\frac{\omega_s}{2H}$, $P_e=(V_s\frac{x_s}{x'_s}E'_q-V_sp\cos(\delta))\sin(\delta)/x_s$, $p=\frac{x_d-x'_d}{x'_s}V_s$, $\sigma=\frac{1}{T'_{d0}}$, $x'_s=x_T+x_L+x'_d$, $x_s=x_T+x_L+x_d$, and the rest parameters with their own physical meanings are the same as those in [30].

Following the transformation method in [30], the SGS in (1) can be represented as

$$\begin{cases}\dot{x}_1&=x_2\\ \dot{x}_2&=\varrho+\mathcal{H}(u-\widetilde{w}\sin(x_1+\delta_0))\end{cases}\tag{2}$$

where $x_1=\delta-\delta_0$, $x_2=\omega$, $\varrho=-\widetilde{\zeta}x_2-\psi(\frac{1}{2}\sin(x_1)\cos(\delta_0)-\sin^2(\frac{x_1}{2})\cos(\delta_0))$, $\widetilde{\zeta}=\frac{D}{2H}+\mathcal{H}K_GT_G$, $\psi=2\mathcal{H}\frac{V_s}{x_s}E_{qs}$, $\mathcal{H}=\frac{\omega_s}{2H}$, $u=T_G\mu$, and $\delta_0$, and $E_{qs}$ denote the steady-state values of the rotor angle, electromagnetic field. $\widetilde{w}=\frac{V_s}{x_s}(E'_q-E_{qs})$ denotes the unknown internal dynamic satisfying $\dot{\widetilde{w}}=-\sigma\widetilde{w}+\gamma\sin^2(\frac{x_1}{2})\cos(\alpha)+\frac{\gamma}{2}\sin(x_1)\sin(\alpha)$, where $\gamma=\sigma p\frac{V_s}{x_s}$, respectively.

In this study, the CIS is assumed to satisfy

$$|u(t)|<\bar{u}\tag{3}$$

where $\bar{u}>0$.

Define $x(t)=\begin{bmatrix}x_1\\x_2\end{bmatrix}$, $f(x(t))=\begin{bmatrix}x_2\\\varrho\end{bmatrix}$, $G=\begin{bmatrix}0\\\mathcal{H}\end{bmatrix}$, $\varpi(x)=-\widetilde{w}\sin(x_1+\delta_0)$, and $\Delta\ell(x(t))=G\varpi(x)$, then the SGS in (2) is modeled by

$$\dot{x}(t)=f(x(t))+Gu(t)+\Delta\ell(x(t)).\tag{4}$$

After taking into account electromagnetic and mechanical features in SGSs, $\varpi(x)$ in $\ell(x(t))$ is assumed to satisfy

$$\|\varpi(x)\|\le\varpi_M(x)\tag{5}$$

where $\varpi_M(x)$ is a known bounded function.

## III. Event-Based Optimal Control of SGSs With CIS

Here, our focus is on developing an optimal control method to tackle the issue of CIS and an ETM to mitigate limited communication resource for the control of SGSs.

### A. Optimal GCC of SGSs With CIS

For convenience of optimal GCC design, we first consider the following nominal SGS by letting $\Delta\ell(x(t))$ equal to 0 in (4)

$$\dot{x}(t) = f(x(t)) + Gu(t). \tag{6}$$

Addressing the challenge of CIS in SGSs, we introduce the following cost functions that are crucial in the design of optimal GCC:

$$\underline{\mathcal{V}}(x,u) = \int_t^\infty \mathrm{e}^{-\xi(\iota-t)}\mathbf{U}(x(\iota),u(\iota))d\iota \tag{7}$$

$$\mathcal{J}(x,u) = \int_t^\infty \mathrm{e}^{-\xi(\iota-t)}(\zeta\varpi_M^2(x) + \mathbf{U}(x(\iota),u(\iota)))d\iota \tag{8}$$

where $\mathbf{U}(x,u) = \mathscr{Q}(x) + \mathscr{B}_u(u)$ with $\mathscr{Q}(x) = \{x|Q\}$ and

$$\mathscr{B}_u(u) = 2\bar{u}\int_0^{u(t)} \operatorname{arctanh}\left(\frac{v}{\bar{u}}\right)dv. \tag{9}$$

In (7) and (8), $\xi > 0$ is a discount factor, $\zeta \geq 1$, and matrix $Q > 0$. In subsequent discourse, for simpleness's sake, a tanh is a shorthand for the function arctanh.

*Theorem 1:* The SGS (4) subject to CIS in (3) is uniformly ultimately bounded (UUB) if there exists a functional $\mathcal{V}(x) > 0$, which is continuously differentiable that satisfies

$$(\nabla\mathcal{V}(x))^T G = -2\bar{u}\operatorname{atanh}(u(x)/\bar{u}) \tag{10}$$

$$\zeta\varpi_M^2(x) + \mathbf{U}(x,u) - \xi\mathcal{V}(x) + (\nabla\mathcal{V}(x))^T(f(x)+Gu(x)) = 0 \tag{11}$$

where $\mathcal{V}(x) = 0$ only if $x = 0$.

*Proof:* Consider the Lyapunov functional (L-functional) $\mathcal{V}(x) > 0$. Inspired by [7], and the $u(x)$ is admissible, then we assume: $|\mathcal{V}(x)| \leq \Lambda_M$ and $\|\nabla\mathcal{V}(x)\| \leq \Theta_M$, where $\Lambda_M$ and $\Theta_M$ are positive constant.

Combining (10) and (11), and deriving L-functional $\mathcal{V}(x)$ along (4) yield that

$$\dot{\mathcal{V}}(x) = -\zeta\varpi_M^2(x) - \{x|Q\} - \mathscr{B}_u(u) + \xi\mathcal{V}(x) - 2\bar{u}\operatorname{atanh}(u(x)/\bar{u})\varpi(x). \tag{12}$$

Utilizing integration by parts method, the Lagrange mean value theorem to (9), and altering (12) involves the addition and subtraction of $\varpi(x)^T\varpi(x)$, for $m \in (0, \operatorname{atanh}(u/k))$, we have

$$\begin{aligned}\dot{\mathcal{V}}(x) \leq\ & -\zeta\varpi_M^2(x) - \{x|Q\} + \xi\mathcal{V}(x) \\ & + 2\bar{u}^2\operatorname{atanh}(u(x)/\bar{u})(m\tanh^2(m)) + \varpi(x)^T\varpi(x) \\ & - \|\bar{u}\operatorname{atanh}(u(x)/\bar{u}) + \varpi(x)\|^2.\end{aligned} \tag{13}$$

Then, recalling (5) and (13) yield that

$$\begin{aligned}\dot{\mathcal{V}}(x) \leq\ & -(\zeta-1)\varpi_M^2(x) - \{x|Q\} + \frac{1}{2}\|G\|^2\Theta_M^2 + \xi\Lambda_M \\ & - \|\bar{u}\operatorname{atanh}(u(x)/\bar{u}) + \varpi(x)\|^2 \\ \leq\ & -\lambda_{\min}(Q)\|x\|^2 + \frac{1}{2}\|G\|^2\Theta_M^2 + \xi\Lambda_M.\end{aligned}$$

Accordingly, $\dot{\mathcal{V}}(x) < 0$ when

$$\|x\| > \sqrt{\frac{\|G\|^2\Theta_M^2 + 2\xi\Lambda_M}{2\lambda_{\min}(Q)}}. \tag{14}$$

Thus, one can conclude that the SGS (4) subject to CIS is UUB. ■

Subsequently, we derive $\mathcal{V}(x)$ along system (6), and it deduces

$$\dot{\mathcal{V}}(x) = (\nabla\mathcal{V}(x))^T(f(x)+Gu(x)). \tag{15}$$

From (11), it has

$$\dot{\mathcal{V}}(x) = \zeta\varpi_M^2(x) + \mathbf{U}(x,u) + \xi\mathcal{V}(x). \tag{16}$$

It can be derived that $\mathcal{J}(x,u)$ in (8) is equivalent to

$$\mathcal{J}(x,u) = -\mathrm{e}^{-\xi(\iota-t)}\mathcal{V}(x)|_{\iota=t}^{\infty} = \mathcal{V}(x). \tag{17}$$

Then, comparing (7) with (8) yields that $\mathcal{V}(x) = \mathcal{J}(x,u) \geq \underline{\mathcal{V}}(x,u)$.

Now, it is asserted that the analysis of the nominal SGS (6) with the cost function (8) using the controller with CIS in (3) corresponds to the optimal GCC for the SGS in (4) with function (7), and $\mathcal{J}(x,u)$ is bounded below by $\underline{\mathcal{V}}(x,u)$.

Following this, our objective is to develop the control law of optimal GCC for SGSs. With the aforementioned analysis, we ought to turn to achieve the minimalization of the lower bounded cost function $\mathcal{J}(x,u)$, then the optimal control law can be obtained. Thus, the nominal SGSs (6) with the cost function (8) should be considered. From the cost function in (8) and (17), it has

$$\mathcal{V}(x) = \int_t^T \mathrm{e}^{-\xi(\iota-t)}(\zeta\varpi_M^2(x) + \mathbf{U}(x,u))d\iota + \mathcal{V}(x(T)). \tag{18}$$

Deriving (18), we obtain

$$\zeta\varpi_M^2(x) + \mathbf{U}(x,u) - \xi\mathcal{V}(x) + \nabla\mathcal{V}^T(x)(f(x)+Gu(x)) = 0. \tag{19}$$

Applying the classic optimal control theory, we define the following Hamiltonian function:

$$\mathrm{H}(x,u,\nabla\mathcal{V}) = \mathbf{U}(x,u) + \zeta\varpi_M^2(x) - \xi\mathcal{V}(x) + (\nabla\mathcal{V}(x))(f(x)+Gu(x)). \tag{20}$$

According to optimality principle proposed by Bellman, the optimal cost function of the nominal SGS in (6) is $\mathcal{V}^*(x) = \min_{u\in\mathscr{A}(\Omega)} \mathcal{V}(x,u)$, and corresponding Hamilton–Jacobi–Bellman (HJB) is formulated by

$$0 = \min_{u\in\mathscr{A}(\Omega)} \mathrm{H}(x,u,\nabla\mathcal{V}^*(x))$$

$$= \{x|Q\} + \mathscr{B}_u(u^*) + \zeta\varpi_M^2(x) - \xi\mathcal{V}^*(x) + (\nabla\mathcal{V}^*(x))^T(f(x) + Gu^*(x)) \tag{21}$$

with $\mathcal{V}^*(0) = 0$.

From (21), one can get the control law of optimal GCC as

$$u^*(x) = -\bar{u}\tanh(\frac{1}{2\bar{u}}G^T\nabla\mathcal{V}^*(x)). \tag{22}$$

Combining (21) with (22) yields

$$0 = \{x|Q\} + \mathscr{B}_u(u^*) + \zeta\varpi_M^2(x) + (\nabla\mathcal{V}^*(x))^T f(x) - \xi\mathcal{V}^*(x) - (\nabla\mathcal{V}^*(x))^T G\bar{u}\tanh\left(\frac{1}{2\bar{u}}G^T\nabla\mathcal{V}^*(x)\right) \tag{23}$$

with $\mathcal{V}^*(0) = 0$.

Recalling (17), it is known that $\mathcal{J}(x, u^*) = \mathcal{V}^*(x)$ by letting $u(x) = u^*(x)$, indicating that $u^*(x) = \arg\min_{u(x)}\mathcal{J}(x, u)$. Consequently, the control law of optimal GCC for the SGS in (4) can be derived from the solution of the continuously differentiable function $\mathcal{V}^*(x)$ in (23) associated with the nominal SGS in (6).

To decrease the frequency of control input updates, the utilization of an event-triggered mechanism becomes essential to decrease the computation resource. Subsequently, an event-triggered GCC strategy will be formulated, building upon the previously discussed optimal GCC strategy in Section III-A.

### B. Event-Based Optimal GCC for SGSs

For convenience of description, we define $\{s_j\}_{j=0}^{\infty}$ with $s_0 = 0$ and $j \in \mathbb{N}$ as triggering instants; and $x_j$ denotes $x(s_j)$. With the aid of ZOH, we have $u(x) = u(x_j)$ for $t \in [s_j, s_{j+1})$.

Then, the nominal SGS in (6) can be rewritten as

$$\dot{x} = f(x) + Gu(x_j) \tag{24}$$

and the control law of optimal GCC in (22) turns to be

$$u^*(x_j) = -\bar{u}\tanh\left(\frac{1}{2\bar{u}}G^T\nabla\mathcal{V}^*(x_j)\right). \tag{25}$$

The following reasonable assumption is necessary for designing the event-triggered mechanism, which is common for the RL-based approach and important to ensure the stability of SGSs.

*Assumption 1:* There exist a vector $\Xi$ such that $\|u^*(x) - u^*(x_j)\|^2 \le \|\Xi(x(s_j) - x(t))\|^2$.

The HJB equation corresponding to (20) is

$$\begin{aligned}&\mathrm{H}(x, u^*(x_j), \nabla\mathcal{V}^*(x))\\ &= \{x|Q\} + \mathscr{B}_u(u^*(x_j)) + \zeta\varpi_M^2(x) + (\nabla\mathcal{V}^*(x))^T f(x)\\ &\quad - \xi\mathcal{V}^*(x) - (\nabla\mathcal{V}^*(x))^T G\bar{u}\tanh\left(\frac{1}{2\bar{u}}G^T\nabla\mathcal{V}^*(x_j)\right)\end{aligned}$$

with $\mathcal{V}^*(0) = 0$.

The event-triggered mechanism is developed as follows:

$$t_{k+1} = \min_{t>t_k}\{t|\mathscr{T}(t) > 0\} \tag{26}$$

where $\mathscr{T}(t) = -\vartheta\{x|Q\} + \{e_j|\Psi\} - \varphi$ with $0 < \vartheta < 1$ and $\varphi > 0$, $e_j(t) = x(s_j) - x(t)$ for $t \in [s_j, s_{j+1})$, and $\Psi = \frac{1}{2}\|\Xi\|^2\|G\|^2$.

*Remark 1:* It is noted that the existing results on the ETM for linear systems employ relative error for threshold design, which is more reasonable and fit for practical applications. Unlike the existing ETMs in nonlinear systems with absolute error for threshold design, such as those in [10] and[21], relative errors are designed for the threshold in the ETM (26), thereby unifying the ETM for linear and nonlinear systems.

*Theorem 2:* Under the ETM in (26), the SGS (4) with the saturated optimal control law $u^*(x_j)$ in (25) is UUB.

Furthermore, the event-based optimal guaranteed cost satisfies

$$\begin{aligned}&\underline{\mathcal{V}}(x, u^*(x_j))\\ &\le \mathcal{V}^*(x) + \int_t^{\infty}\mathrm{e}^{-\xi(\iota-t)}\int_{u^*(x(\iota))}^{u^*(x_j)}2\bar{u}\mathrm{atan}(\tau/\bar{u})d\tau d\iota\\ &\quad + \int_t^{\infty}\mathrm{e}^{-\xi(\iota-t)}(\nabla\mathcal{V}^*(x(\iota)))^T G(u^*(x_j) - u^*(x(\iota)))d\iota\\ &\quad + \int_t^{\infty}\mathrm{e}^{-\xi(\iota-t)}\bar{u}^2(\mathrm{atan}(u^*(x(\iota))/\bar{u}))^2 d\iota.\end{aligned} \tag{27}$$

Particularly, the time-based optimal guaranteed cost satisfies

$$\underline{\mathcal{V}}(x, u^*(x)) \le \mathcal{V}^*(x) + \int_t^{\infty}\mathrm{e}^{-\xi(\iota-t)}\bar{u}^2(\mathrm{atan}(u^*(x(\iota))/\bar{u}))^2 d\iota. \tag{28}$$

*Proof:* Under the saturated optimal control law $u^*(x_j)$ in (25), deriving $\mathcal{V}^*(x)$ along the SGS (4), we obtain

$$\dot{\mathcal{V}}^*(x) = (\nabla\mathcal{V}^*(x))^T(f(x) + Gu^*(x_j) + \Delta\ell(x)). \tag{29}$$

Taking into account (21), we can derive

$$\begin{aligned}(\nabla\mathcal{V}^*(x))^T f(x) = &-\{x|Q\} - \mathscr{B}_u(u^*(x)) - \zeta\varpi_M^2(x)\\ &+ \xi\mathcal{V}^*(x) - (\nabla\mathcal{V}^*(x))^T Gu^*(x).\end{aligned} \tag{30}$$

From (5), (22), and (30), (29) is reformulated as

$$\begin{aligned}\dot{\mathcal{V}}^*(x) \le &-\{x|Q\} - \|\varpi(x) + \bar{u}\mathrm{atanh}(u^*(x)/\bar{u})\|^2\\ &+ 2\bar{u}^2(\mathrm{atanh}(u^*(x)/\bar{u}))^2 + \xi\mathcal{V}^*(x)\\ &- (\nabla\mathcal{V}^*(x))^T G(u^*(x_j) - u^*(x)).\end{aligned} \tag{31}$$

According to Assumption 1, we have

$$\begin{aligned}&(u^*(x) - u^*(x_j))^T G^T G(u^*(x) - u^*(x_j))\\ &\le \|G\|^2\|(\Xi(x(s_j) - x(t)))\|^2.\end{aligned} \tag{32}$$

Applying Young's inequality, we have

$$\dot{\mathcal{V}}^*(x) \le \mathscr{T}(t) + \varphi_0 + \varphi - (1-\vartheta)\|Q\|\|x\|^2 \tag{33}$$

where $\varphi_0 = \frac{1}{2}(1 + \|G\|^2)\Theta_M^2 + \xi\Lambda_M$.

Recall the ETM in (26), and we knows that $\dot{\mathcal{V}}^*(x) \le 0$ when

$$\|x\|^2 \ge \frac{\varphi + \varphi_0}{(1-\vartheta)\|Q\|}. \tag{34}$$

Therefore, the SGS (4) is UUB under the ETM in (26) and the saturated control law of optimal GCC in (25).

*Remark 2:* Observing the ETM in (26) and (32) knows that the utilization of the ETM will degrade the SGS performance, however, it can greatly save the computation resource. The parameters $\varphi$ and $\vartheta$ in (26) are set to conveniently balance between the control and the computation efficiency.

Subsequently, we shall analysis the optimal guaranteed cost in (7) of the SGS under the proposed ETM in (26) and the saturated optimal control law in (25).

From (7), it has

$$\underline{\mathcal{V}}(x,u)=\mathcal{V}^*(x)+\int_t^\infty[\mathbf{U}(x,u)+\dot{\mathcal{V}}^*(x)]d\tau.$$

Under the admissible control law $u(x)$, take the derivative of $\mathcal{V}^*(x)$ along the trajectory of the SGS (4), then we have

$$\dot{\mathcal{V}}^*(x)=(\nabla\mathcal{V}^*(x))^T(f(x)+Gu(x)+\Delta\ell(x)).$$

From (30), one can obtain

$$\begin{aligned}&\{x|Q\}+\mathscr{B}_u(u(x))\\&\quad=-\zeta\varpi_M^2(x)+(\mathcal{V}^*(x))^T\Delta\ell(x(t))\\&\qquad+(\nabla\mathcal{V}^*(x))^TG(u(x)-u^*(x))\\&\qquad+\xi\mathcal{V}^*(x)-\dot{\mathcal{V}}^*(x)+\mathscr{B}_u(u(x))-\mathscr{B}_u(u^*(x)).\end{aligned}\tag{35}$$

Considering (5), (8), and (12), we have

$$\begin{aligned}&-\zeta\varpi_M^2(x)+(\mathcal{V}^*(x))^T\Delta\ell(x(t))\\&\qquad\le-\|\varpi(x)+\bar{u}\text{atanh}(u^*(x)/\bar{u})\|^2\\&\qquad\quad+\bar{u}^2(\text{atanh}(u^*(x)/\bar{u}))^2.\end{aligned}\tag{36}$$

Combining (35)–(36) follows that

$$\begin{aligned}\{x|Q\}+\mathscr{B}_u(u(x))\le\xi\mathcal{V}^*(x)-\dot{\mathcal{V}}^*(x)\\+\int_{u^*(x)}^{u(x)}2\bar{u}\text{atanh}(\tau/\bar{u})d\tau\\+(\nabla\mathcal{V}^*(x))^TG(u(x)-u^*(x))\\+\bar{u}^2(\text{atanh}(u^*(x)/\bar{u}))^2.\end{aligned}\tag{37}$$

Integrating both sides of (37) over the interval from $t$ to $\infty$ and taking into account the optimal guaranteed cost in (7), one yields inequality (27).

In particularly, if one sets $u^*(x_j)=u^*(x)$ for (27), then the time-based optimal guaranteed cost of the SGS presented in (28) can be achieved. The proof is completed. ■

## IV. RL-Based Algorithm and Stability Analysis

In Section III, the event-triggered optimal GCC of SGSs is designed in Theorem 2. However, the item $\nabla\mathcal{V}^*(x_j)$ in (25) is hard to obtain. This section presents the development of an online RL-based algorithm aimed at addressing this challenge. By training the critic neural network (CNN) shown in Fig. 1, the optimal cost function can be approximated, and the weight vector needed in the controller can be obtained. Applying the optimal control theory, the event-based optimal GCC under the CIS can be implemented such that the SGS is UUB.

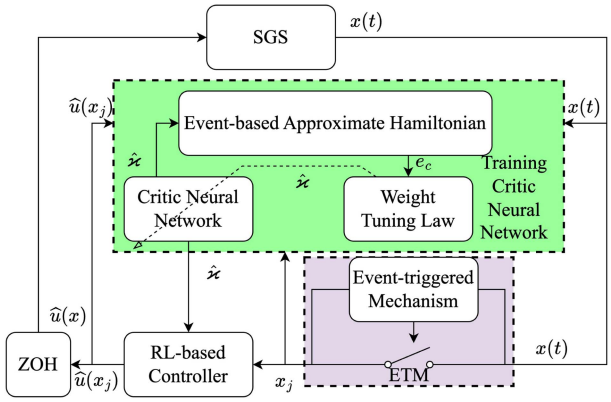


Fig. 1. Control diagram of the RL-based SGS.

### A. RL-Based Iteration Algorithm for Event-Triggered Optimal GCC Implementation

As indicated in Theorems 1, $\mathcal{V}(x)$ is continuously differentiable. Then, it can be reconstructed within a compact set $\Omega$ using the neural network thanks to its universal approximation property as follows:

$$\mathcal{V}^*(x)=\varkappa_c^T\Im_c(x)+\wp_c(x)\tag{38}$$

where $\varkappa_c\in\mathbb{R}^{l_c}$ represents the perfect weights, $\Im_c(x)$ is the activation function, $l_c$ denotes the hidden layer neuron count, and $\wp_c(x)$ stands for the unknown estimated neural network error.

Then, the saturated optimal control law under ETM in (25) can be replaced by

$$u^*(x_j)=-\bar{u}\tanh(\mathscr{G}\varkappa_c+\mathscr{J}_r)\tag{39}$$

where $\mathscr{G}=\frac{1}{2\bar{u}}G^T\nabla\Im_c^T(x_j)$, and $\mathscr{J}_r=\frac{1}{2\bar{u}}G^T\nabla\wp_c(x_j)$.

*Remark 3:* In this study, a single neural network is employed to approximate both the cost function and the controller, effectively minimizing the superfluous consumption of computing resources that would result from constructing another separate network dedicated to the controller.

The HJB equation in (21) can be rewritten as

$$\begin{aligned}\text{H}(x,u^*(x),\varkappa_c)=\{x|Q\}+\mathscr{B}_u(u^*(x))+\zeta\varpi_M^2(x)\\+(\nabla\Im_c(x)^T\varkappa_c+\nabla\wp_c(x))(f(x)+Gu^*(x))\\-\xi(\varkappa_c^T\Im_c(x)+\wp_c(x))=0.\end{aligned}\tag{40}$$

It is a challenge that the weight $\varkappa_c$ in (38) is unknown. To address this, the following CNN is built to estimate the cost function $\mathcal{V}^*(x)$:

$$\hat{\mathcal{V}}(x)=\hat{\varkappa}_c^T\Im_c(x).\tag{41}$$

Accordingly, the event-triggered optimal saturated control law in (39) and the HJB equation are approximated by $\hat{u}(x_j)$

and $\text{Ĥ}(x,\hat{u}(x_j),\hat{\varkappa}_c)$, respectively,

$$\hat{u}(x_j) = -\bar{u}\tanh(\mathscr{G}\hat{\varkappa}_c) \tag{42}$$

$$\begin{aligned}\text{Ĥ}(x,\hat{u}(x_j),\hat{\varkappa}_c) &= \{x|Q\} + \mathscr{B}_u(\hat{u}(x_j)) + \zeta\varpi_M^2(x)\\ &\quad +\nabla\Im_c(x)^T\hat{\varkappa}_c(f(x)+G\hat{u}(x_j)) - \xi\hat{\varkappa}_c^T\Im_c(x).\end{aligned} \tag{43}$$

Accordingly, the closed-loop SGS with the approximated control law of optimal GCC in (42) becomes

$$\dot{x} = f(x) - G\bar{u}\tanh(\frac{1}{2\bar{u}}G^T\nabla\Im_c^T(x_j)\hat{\varkappa}_c). \tag{44}$$

Inspired by [7], the steepest descent algorithm is used to train the CNN, aiming to minimize $E_c$, where $E_c = 0.5\|e_c\|^2$ with $e_c = \text{Ĥ}(x,\hat{u}(x_j),\hat{\varkappa}_c) - \text{Ĥ}(x,u^*(x),\varkappa_c)$. According to (40), $e_c$ becomes $e_c = \text{Ĥ}(x,\hat{u}(x_j),\hat{\varkappa}_c)$. The tuning law of $\hat{\varkappa}_c$ is designed by

$$\dot{\hat{\varkappa}}_c = -\alpha_c\mathscr{X}^2\mathfrak{P}(E_c) - \mathscr{U}(x,u(x_j))\alpha_s\mathfrak{P}(\mathcal{V}_e) \tag{45}$$

where $\alpha_c > 0$ and $\alpha_s > 0$ are learning rates; $\mathcal{V}_e(x)$ is a continuously differentiable L-functional; and

$$\mathfrak{P}(E_c) = \chi(\{x|Q\} + \mathscr{B}_u(\hat{u}(x_j)) + \zeta\varpi_M^2(x) + \chi^T\hat{\varkappa}_c)$$

$$\mathfrak{P}(\mathcal{V}_e) = \mathit{\Pi}\nabla\mathcal{V}_e(x)$$

$$\mathit{\Pi} = 0.5\nabla\Im_c(x_j)G\text{sech}^2\mathscr{Z}G^T$$

$$\mathscr{Z} = \frac{1}{2\bar{u}}G^T\nabla\Im^T(x_j)\hat{\varkappa}_c,\ \mathscr{X} = (1+\chi^T\chi)^{-1}$$

$$\chi = \nabla\Im_c(x)(f(x)+G\hat{u}(x_j)) - \xi\Im_c(x)$$

$$\mathscr{U}(x,\hat{u}(x_j)) = \begin{cases}0, & \nabla\mathcal{V}_e^T(x)(f(x)+G\hat{u}(x_j)) < 0\\ 1, & \text{else}\end{cases}.$$

*Remark 4:* In conventional tuning laws, the admissible control law is required for training weight vector of the CNN. However, in practice, it is hard to obtain an admissible control law. In (45), the incorporation of the sign function $\mathscr{U}(x,u(x_j))$ serves to relax the necessity of the requirement for ACP. When the system is unstable, it will be activated to reinforce training process of the weights to stabilize system. Otherwise, it defaults to the conventional tuning law. In facts, the second item in in (45) is a gradient of $\dot{\mathcal{V}}_e$, due to the fact

$$\frac{\partial(\dot{\mathcal{V}}_e)}{\partial\hat{\varkappa}_c} = \frac{\partial\mathscr{Z}}{\partial\hat{\varkappa}_c}\frac{\partial(\bar{u}\nabla\mathcal{V}_e^T(x)G\tanh(\mathscr{Z}))}{\partial\mathscr{Z}} = \mathfrak{P}(\mathcal{V}_e).$$

Algorithm 1 clearly showcases the whole RL-based iteration progress of the weight vector $\hat{\kappa}$.

**Algorithm 1:** RL-Based Iteration Algorithm of the Weight Vector.

**Initialization:** Initialize parameters of the ETM: $H$, $\varphi$, $\vartheta$; the cost function: $Q$, $\xi$, $\zeta$; the learning rates: $\alpha_c$ and $\alpha_s$, and the computation threshold: $\emptyset$.
**Output:** weight vector $\hat{\varkappa}_c$ of CNN.
**while** $\hat{\mathcal{V}}(x)_{i+1} - \hat{\mathcal{V}}(x)_i < \emptyset$ **do**
  $e_j = x(s_j) - x(t)$;
  $\mathscr{T}(t) = -\vartheta\{x|Q\} + \{e_j|\Psi\} - \varphi$;
  **if** $\mathscr{T}(t) > 0$ **then**
    $x(s_j) = x$;
    $\hat{u}(x_j) = -\bar{u}\tanh(\mathscr{G}\hat{\varkappa}_c)$;
  **else**
    $x(s_j) = x(s_{j-1})$;
    $\hat{u}(x_j) = \hat{u}(x_{j-1})$;
  **end if**
  $e_c = \text{Ĥ}(x,\hat{u}(x_j),\hat{\varkappa}_c)$;
  $\dot{\hat{\varkappa}}_c = -\alpha_c\mathscr{X}^2\mathfrak{P}(E_c) - \mathscr{U}(x,u(x_j))\alpha_s\mathfrak{P}(\mathcal{V}_e)$;
  $i = i+1$;
**end while**

## B. Convergence Analysis of the RL-Based SGSs

This section will focus on analyzing the stability of the SGS and the convergence of the RL-based algorithm.

Defining $\tilde{\varkappa}_c = \varkappa_c - \hat{\varkappa}_c$, it follows that

$$\begin{aligned}\dot{\tilde{\varkappa}}_c = &-\alpha_c\mathscr{X}^2\chi(\chi^T\tilde{\varkappa}_c - \epsilon_{cH})\\ &-\frac{1}{2}\mathscr{U}(x,u(x_j))\alpha_s\mathit{\Pi}\nabla\mathcal{V}_e(x)\end{aligned} \tag{46}$$

where $\epsilon_{cH} = (\nabla\wp_c(x))^T(f(x)+G\hat{u}(x_j)) + \xi\wp_c(x)$, which is assumed to satisfy $\|\epsilon_{cH}\| \le \lambda_e$.

Inspired by [4], we do the following assumption.

*Assumption 2:* There exists a positive definite matrix $\Lambda(x)$ such that $(\nabla\mathcal{V}_e(x))^T(f(x)+Gu^*(x_j)) = -(\nabla\mathcal{V}_e(x))^T\Lambda(x)(\nabla\mathcal{V}_e(x))$.

*Theorem 3:* Consider the SGS in (6) with the approximate optimal saturated control law in (42), the ETM in (26), and the tuning law in (45). UUB of the closed-loop SGS in (44) and the weight error system are guaranteed.

*Proof:* L-functional is constructed as follows:

$$\mathfrak{T}(t) = \mathfrak{T}_1(t) + \mathfrak{T}_2(t) + \mathfrak{T}_3(t) \tag{47}$$

where $\mathfrak{T}_1(t) = \frac{1}{2\alpha_c}\tilde{\varkappa}_c^T\tilde{\varkappa}_c + \frac{\alpha_s}{\alpha_c}\mathcal{V}_e(x)$, $\mathfrak{T}_2(t) = \mathcal{V}^*(t)$, and $\mathfrak{T}_3(t) = \mathcal{V}^*(x_j)$.

To analysis the stabilities of both the closed-loop SGS in (44) and the weigh error system of CNN (46). Two cases will be discussed in the following due to the discontinuity of $\mathfrak{T}_3(t)$: when $t$ belongs to the interval $[s_j, s_{j+1})$ and when $t$ is equal to $s_j$.

*Case 1*: We consider the case that $t \in [s_j, s_{j+1})$. The state in this interval maintains the value at $s_j$, i.e. the event is not triggered.

In (45), the sign function $\mathscr{U}(x,u(x_j))$ is introduced to expedite the convergence speed of neural network training. For $\mathscr{U}(x,u(x_j)) = 1$, we have the following derivation.

Taking the derivative of $\mathfrak{T}_1(t)$ and considering (39) lead to

$$\begin{aligned}\dot{\mathfrak{T}}_1(t) &= \frac{1}{\alpha_c}\tilde{\varkappa}_c^T\dot{\tilde{\varkappa}}_c + \frac{\alpha_s}{\alpha_c}\dot{\mathcal{V}}_e(x)\\ &= \hbar(x) + \bar{u}\mathscr{H}(\mathscr{D}_1 - \mathscr{D}_2 - \text{sech}^2(\mathscr{G}\hat{\kappa})\mathscr{G}\tilde{\kappa})\end{aligned} \tag{48}$$

where $\hbar(x) = -\tilde{\varkappa}_c^T\chi\mathscr{X}^2(\chi^T\tilde{\varkappa}_c - \epsilon_{cH}) + \frac{\alpha_s}{\alpha_c}(\nabla\mathcal{V}_e(x))^Tf(x)$, $\mathscr{H} = \frac{\alpha_s}{\alpha_c}(\nabla V_e^T(x)G$, $\mathscr{D}_1 = \tanh(\mathscr{G}\varkappa) - \tanh(\mathscr{G}\hat{\kappa})$, and $\mathscr{D}_2 = \tanh(\mathscr{G}\kappa)$.

Using the Taylor expansion to $\mathscr{D}_1$ and $\mathscr{D}_2$, we have

$$\dot{\mathfrak{T}}_1(t) = \hbar(x) + \mathscr{H}[u^*(x_j) - \wp_{u^*} + \bar{u}\mathscr{O}((\mathscr{G}\tilde{\kappa})^2)] \tag{49}$$

where $\mathscr{O}(\cdot)$ is higher order term of the Taylor series and $\wp_{u^*} = \bar{u}(1 - \frac{1}{\bar{u}^2}u^*(x_j))(-\mathscr{J}_r) - \bar{u}\mathscr{O}((-\mathscr{J}_r)^2)$, from which we can observe that $\|\wp_{u^*}\| \leq \lambda_\wp$.

Observing the fact that higher order term $\mathscr{O}((\mathscr{G}\tilde{\kappa})^2)$ is bounded by for hyperbolic function tanh

$$\|\mathscr{O}\left((\mathscr{G}\tilde{\kappa})^2\right)\| \leq \mathscr{C}_1 + \mathscr{C}_2\|\tilde{\varkappa}_c\| \tag{50}$$

where $\mathscr{C}_i (i = 1, 2)$ are positive constants.

Applying Young's inequality and noticing (50) together with $1 + \chi^T\chi \geq 1$, $\|\epsilon_{cH}\| \leq \lambda_e$ and $\|\wp_{u^*}\| \leq \lambda_\wp$, we have

$$\begin{aligned}\dot{\mathfrak{T}}_1(t) \leq & -(\lambda_0 - \bar{u}\mathscr{C}_2\lambda_1^2\lambda_3/\lambda_2)\|\tilde{\varkappa}_c\|^2 + \frac{1}{2}\lambda_e^2 \\ & - \beta\lambda_2\|\nabla\mathcal{V}_e(x)\|^2 + \lambda_1\lambda_4\|\nabla\mathcal{V}_e(x)\|\end{aligned}$$

where $\lambda_0 = \frac{1}{2}\lambda_{\min}(\chi\chi^T)\mathscr{X}^2$, $\lambda_1 = \frac{\alpha_s}{\alpha_c}\|G\|$, $\lambda_2 = \frac{\alpha_s}{\alpha_c}\lambda_{\min}(\Lambda(x))$, $\lambda_3 = 0.25(1-\beta)^{-1}$, $\lambda_4 = \lambda_\wp + \bar{u}\mathscr{C}_1$, and $\beta \in (0, 1)$.

Then, we can obtain $\dot{\mathfrak{T}}_1(t) \leq 0$ when

$$\|\tilde{\varkappa}\|^2 \geq \frac{\beta\lambda_2\lambda_e^2 - \lambda_1\lambda_4 + \lambda_1^3(\lambda_\wp + \bar{u}\mathscr{C}_1^2)}{2\beta\lambda_0\lambda_2 - 2\bar{u}\mathscr{C}_2\lambda_1^2\lambda_3} \tag{51}$$

or

$$\|\nabla\mathcal{V}_e(x)\| \geq \frac{\lambda_1\lambda_4 + \sqrt{\lambda_1^2\lambda_4^2 + 2\beta\lambda_2\lambda_e}}{2\beta\lambda_2} \tag{52}$$

holds.

From (33), one can obtain

$$\dot{\mathfrak{T}}_2(t) + \dot{\mathfrak{T}}_3(t) \leq \mathscr{T}(t) + \varphi_0 + \varphi - (1-\vartheta)\|Q\|\|x\|^2. \tag{53}$$

Thus, when states of the SGS violates condition (34)

$$\dot{\mathfrak{T}}_{23}(t) = \dot{\mathfrak{T}}_2(t) + \dot{\mathfrak{T}}_3(t) \leq 0 \tag{54}$$

and the SGS under the ETM in (26) is UUB.

For $\mathscr{U}(x, u(x_j)) = 0$, similar to [4], we have $0 < a\|\nabla\mathcal{V}_e(x)\| < -(\nabla\mathcal{V}_e(x))^T(f(x) + Gu(x_j))$ for $a > 0$. Hence, we have

$$\dot{\mathfrak{T}}_1 \leq \frac{1}{2}\|\epsilon_{cH}\|^2 - \lambda_0\|\tilde{\varkappa}_c\|^2 - a\|\nabla\mathcal{V}_e(x)\|$$

$$\dot{\mathfrak{T}}_{23} \leq \mathscr{T}(t) + \varphi_0 + \varphi - (1-\vartheta)\|Q\|\|x\|^2.$$

Therefore, given that

$$\|\tilde{\varkappa}\|^2 \geq \frac{\|\epsilon_{cH}\|^2}{2\lambda_0} \text{ or } \|\nabla\mathcal{V}_e(x)\| \geq \frac{\|\epsilon_{cH}\|^2}{2a}$$

holds, it yields $\dot{\mathfrak{T}}_1(t) \leq 0$.

Under EMT (26), we have $\dot{\mathfrak{T}}_{23}(t) \leq 0$ when state violates condition (34).

Considering $\dot{\mathfrak{T}}_1(t) \leq 0$ and $\dot{\mathfrak{T}}_{23}(t) \leq 0$, the UUB of the systems (44) and (46) are obtained when $\mathscr{U}(x, u(x_j)) = 0$.

*Case 2:* We consider the case when $t = s_j$. In this case, the proof for $\mathscr{U}(x, u(x_j)) = 0$ is similar to that when $\mathscr{U}(x, u(x_j)) = 1$. Therefore, our main focus will be on the latter, outlined as follows:

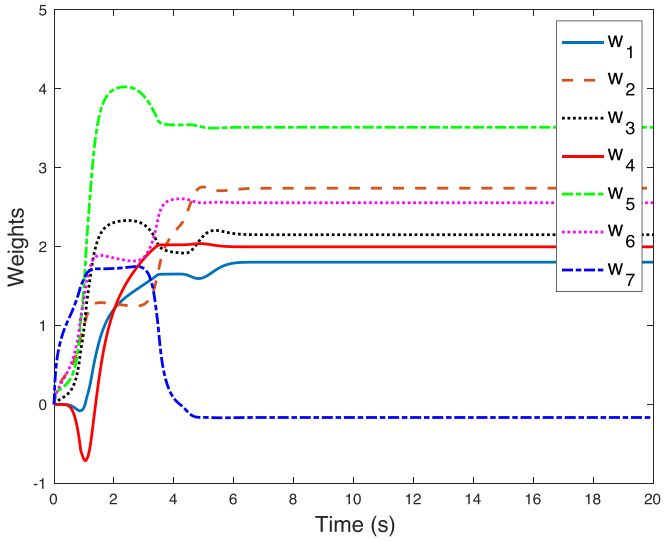


Fig. 2. Convergence process of the weights.

For the discontinuity of $\mathfrak{T}_3(t)$ at $s_j$, we define $\Delta\mathfrak{T}_i$ as the forward difference. Then, it has

$$\Delta\mathfrak{T}(t) = \Delta\mathfrak{T}_1(t) + \Delta\mathfrak{T}_2(t) + \Delta\mathfrak{T}_3(t).$$

In Case 1, we have proved $\dot{\mathfrak{T}}_i < 0, i = 1, 2$ for all $t \in [s_j, s_{j+1})$. It follows that

$$\Delta\mathfrak{T}_i(t) = \mathfrak{T}_1(\lim_{\delta\to 0}(t+\delta)) - \mathfrak{T}_i(t) \leq 0,\ i = 1, 2. \tag{55}$$

Inspired by [31], we have

$$\Delta\mathfrak{T}_3(t) = \mathfrak{T}_3(s_{j+1}) - \mathfrak{T}_3(s_j) \leq \varsigma(s_{j+1} - s_j)$$

where $\varsigma(\cdot)$ denotes class-$\kappa$ functions.

Based on the aforementioned analysis, systems (44) and (46) are UUB. ■

## V. Simulation Results

The SGS presented in Section II with the parameters $\delta_0$, $K_G$, $V_s$, $x'_d$, $x_d$, $x_T$, $x_L$, $D$, $T_G$, and $T'_{d0}$ set to 1.000, 1.000, 1.000, 0.257, 1.836, 0.127, 0.485, 5.000, 0.200, and 2.000 [30], respectively, is considered to verify the effectiveness of the proposed event-triggered RL-based optimal GCC.

The initial state of the SGS is assumed to be $x_0 = [0.8, 2]^T$, and the control limitation is $\bar{u} = 2.5(\text{s}\cdot\text{r/min/Hz})$. The parameters in the cost function are selected as follows: $\{x|Q\} = 100x_1^2 + 10x_2^2, \zeta = 1.1, \xi = 0.6$, and $\varpi_M(x) = |\widetilde{w}|$. The activation function of the CNN is $[x_1^2, x_2^2, x_1x_2, x_1^3, x_1^2x_2, x_1x_2^2, x_2^3]^T$, and the weights are given by $\hat{\varkappa} = [\hat{\varkappa}_1, \hat{\varkappa}_2, \hat{\varkappa}_3, \hat{\varkappa}_4, \hat{\varkappa}_5, \hat{\varkappa}_6, \hat{\varkappa}_7]^T$. The learning rates are $\alpha_c = 1.9$ and $\alpha_s = 1.5$, respectively. The initial weights are all set as zero. The parameters in the ETM (26) are chosen as $\Xi = [-10, 100]$, and $\varphi = 0.0001$.

Two simulation steps are outlined: initial training phase to determine CNN weights, followed by performance testing using these weights. In the training phase, a probing noise, $N = 0.02(\sin^2(t)\cos(t) + \sin^2(2t)\cos(0.1t) + \sin^2(-1.2t)\cos(0.5t) + \sin^5(t) + \sin^2(1.12t) + \cos(2.4t)\sin^3(2.4t))$, is

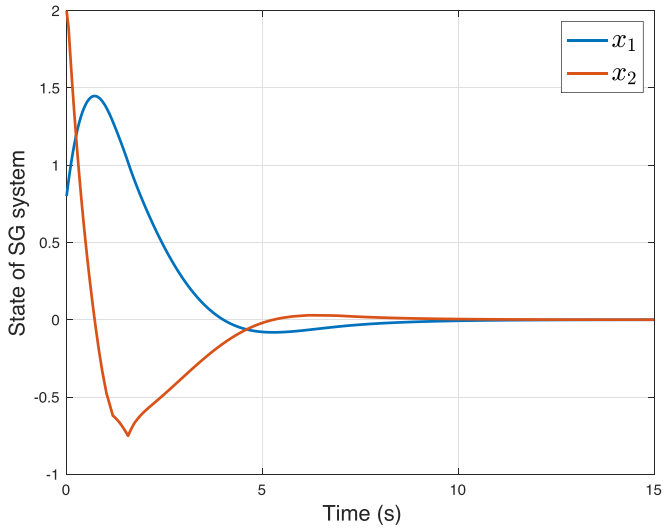


Fig. 3. Trajectory of event-based states.

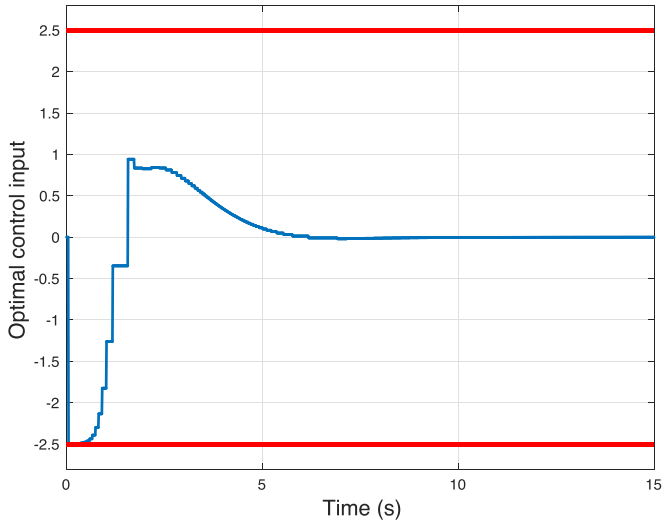


Fig. 4. Trajectory of the control input.

introduced in the control input to obtain the persistence of excitation. Fig. 2 shows the trajectories of the CNN weights, indicating convergence to their final values of $[1.8020, 2.7411, 2.1498, 1.9965, 3.5117, 2.5550, -0.1665]^T$ after around 7 s.

In the second phase, the trained CNN is utilized for stabilizing the SGS. Under the ETM described in (26) and the RL-based control law of optimal GCC in (25), the state trajectories of the SGS are depicted in Fig. 3. It is noteworthy that some items are modeled as dynamic internal disturbance $\widetilde{w}$, which poses challenges for linear control methods. However, this strategy effectively addresses this issue. Fig. 4 illustrates the optimal saturated control input, indicating that the constraints specified in (3) are satisfactorily ensured by employing the RL-based control law of optimal GCC as stated in (25). The curves in Fig. 3 clearly show that the proposed event-triggered optimal GCC, in conjunction with the trained CNN, yields favorable control performance. In addition, the designed ETM optimizes computation

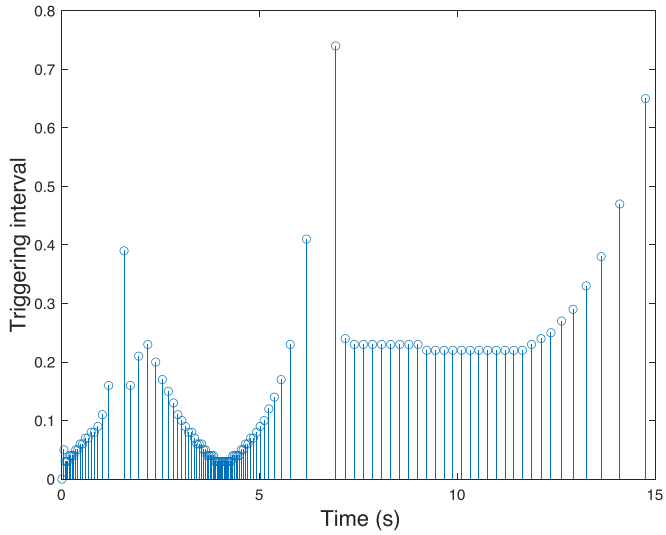


Fig. 5. Triggering interval and triggering instant.

resources, as depicted in Fig. 5. The average triggering period, 0.1353 s, significantly exceeds the nominal sampling period of 0.01 s, demonstrating the efficacy of the ETM in alleviating computational demands. This underscores the pivotal role of the ETM for controlling SGSs efficiently.

## VI. Conclusion

In this article, an RL-based control strategy was investigated for SGSs subject to CIS. The optimal GCC along with an ETM was derived within the RL framework. It was evident from the simulation results that the proposed ETM can extremely alleviate the burden on computation resources, which was crucial for SGSs, especially those employing RL-based control systems that demand complex calculation in each control step. Furthermore, the control performance of SGSs subject to CIS was ensured by the proposed RL-based optimal GCC. Future investigations will focus on RL-based SGSs under varying loads, a more representative scenario for industrial applications.

## References

[1] P. Moriarty and D. Honnery, "What is the global potential for renewable energy?," *Renewable Sustain. Energy Rev.*, vol. 16, no. 1, pp. 244–252, 2012.

[2] Y. Tang, H. He, Z. Ni, J. Wen, and X. Sui, "Reactive power control of grid-connected wind farm based on adaptive dynamic programming," *Neurocomputing*, vol. 125, pp. 125–133, Feb. 2014.

[3] P. S. Kundur and O. P. Malik, *Power System Stability and Control*. New York, NY, USA: McGraw-Hill, 2022.

[4] D. Liu, X. Yang, D. Wang, and Q. Wei, "Reinforcement-learning-based robust controller design for continuous-time uncertain nonlinear systems subject to input constraints," *IEEE Trans. Cybern.*, vol. 45, no. 7, pp. 1372–1385, Jul. 2015.

[5] C. Mu, K. Wang, Z. Ni, and C. Sun, "Cooperative differential game-based optimal control and its application to power systems," *IEEE Trans. Ind. Inform.*, vol. 16, no. 8, pp. 5169–5179, Aug. 2020.

[6] Z.-Q Liu, X. Ge, Q.-L. Han, and Y.-L. Wang, "Reinforcement learning-based cooperative target tracking control of unmanned surface vehicles under data falsification attacks," *IEEE Trans. Intell. Veh.*, early access, May 2024, doi: 10.1109/TIV.2024.3396171.

[7] K. G. Vamvoudakis and F. L. Lewis, "Online actor–critic algorithm to solve the continuous-time infinite horizon optimal control problem," *Automatica*, vol. 46, no. 5, pp. 878–888, 2010.
[8] O. Qasem, W. Gao, and K. G. Vamvoudakis, "Adaptive optimal control of continuous-time nonlinear affine systems via hybrid iteration," *Automatica*, vol. 157, Nov. 2023, Art. no. 111261.
[9] M. Shen, X. Wang, S. Zhu, and T. Huang, "Data-driven event-triggered adaptive dynamic programming control for nonlinear systems with input saturation," *IEEE Trans. Cybern.*, vol. 54, no. 2, pp. 1178–1188, Feb. 2024.
[10] R. Song and L. Liu, "Event-triggered constrained robust control for partly-unknown nonlinear systems via ADP," *Neurocomputing*, vol. 404, pp. 294–303, Sep. 2020.
[11] L. Yu, J. Lai, J. Xiong, and M. Xie, "Robust ADP-based control for uncertain nonlinear Stackelberg games," *Neurocomputing*, vol. 561, Dec. 2023, Art. no. 126834.
[12] R. Song, F. L. Lewis, and Q. Wei, "Off-policy integral reinforcement learning method to solve nonlinear continuous-time multiplayer nonzero-sum games," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 28, no. 3, pp. 704–713, Mar. 2017.
[13] J. Li, J. Ding, T. Chai, F. L. Lewis, and S. Jagannathan, "Adaptive interleaved reinforcement learning: Robust stability of affine nonlinear systems with unknown uncertainty," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 33, no. 1, pp. 270–280, Jan. 2022.
[14] Z. Gu, X. Huang, X. Sun, X. Xie, and J. H. Park, "Memory-event-triggered tracking control for intelligent vehicle transportation systems: A leader-following approach," *IEEE Trans. Intell. Transp. Syst.*, vol 25, no. 5, pp. 4021–4031, May 2024.
[15] D. Yue, E. Tian, and Q.-L. Han, "A delay system method for designing event-triggered controllers of networked control systems," *IEEE Trans. Autom. Control*, vol. 58, no. 2, pp. 475–481, Feb. 2013.
[16] B. C. Zheng, X. H. Yu, and Y. M. Xue, "Quantized feedback sliding-mode control: An event-triggered approach," *Automatica*, vol. 91, pp. 126–135, May 2018.
[17] K. Wang and C. Mu, "Learning-based control with decentralized dynamic event-triggering for vehicle systems," *IEEE Trans. Ind. Inform.*, vol. 19, no. 3, pp. 2629–2639, Mar. 2023.
[18] S. Hu, D. Yue, Z. Cheng, E. Tian, X. Xie, and X. Chen, "Co-design of dynamic event-triggered communication scheme and resilient observer-based control under aperiodic DoS attacks," *IEEE Trans. Cybern.*, vol. 51, no. 9, pp. 4591–4601, Sep. 2021.
[19] C. Mu, K. Wang, and C. Sun, "Learning control supported by dynamic event communication applying to industrial systems," *IEEE Trans. Ind. Inform.*, vol. 17, no. 4, pp. 2325–2335, Apr. 2021.
[20] Y. Shi, Q. Hu, D. Li, and M. Lv, "Adaptive optimal tracking control for spacecraft formation flying with event-triggered input," *IEEE Trans. Ind. Inform.*, vol. 19, no. 5. pp. 6418–6428, May 2023.
[21] 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.
[22] Z. Gu, Y. Fan, X. Sun, X. Xie, and C. K. Ahn, "Event-based two-step transmission mechanism for the stabilization of networked T–S fuzzy systems with random uncertainties," *IEEE Trans. Cybern.*, vol. 54, no. 2, pp. 1283–1293, Feb. 2024.
[23] A. E. Hoffer, R. H. Moncada, B. J. Pavez-Lazo, J. A. Tapia, and L. Laurila, "Calculation of a current vector trajectory for enhanced operation of synchronous reluctance generators including saturation," *IEEE Trans. Ind. Electron.*, vol. 70, no. 2, pp. 1197–1204, Feb. 2023.
[24] A. Leon, J. Solsona, J. Figueroa, and M. Valla, "Optimization with constraints for excitation control in synchronous generators," *Energy*, vol. 36, no. 8, pp. 5366–5373, Aug. 2011.
[25] X. Yang, Y. Zhou, and Z. Gao, "Reinforcement learning for robust stabilization of nonlinear systems with asymmetric saturating actuators," *Neural Netw.*, vol. 158, pp. 132–141, Jan. 2023.
[26] J. Qin, M. Li, Y. Shi, Q. Ma, and W. X. Zheng, "Optimal synchronization control of multiagent systems with input saturation via off-policy reinforcement learning," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 30, no. 1, pp. 85–96, Jan. 2019.
[27] H. Zhan, D. Huang, and C. Yang, "Adaptive dynamic programming enhanced admittance control for robots with environment interaction and actuator saturation," *Int. J. Intell. Robot. Appl.*, vol. 5, no. 1, pp. 89–100, Feb. 2021.
[28] 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. Autom. Sci. Eng.*, vol. 21, no. 3, pp. 4674–4685, Jul. 2024, doi: 10.1109/TASE.2023.3300917.
[29] Y. Zhang, Q. Sun, J. Zhou, J. M. Guerrero, R. Wang, and A. Lashab, "Optimal frequency control for virtual synchronous generator based ac microgrids via adaptive dynamic programming," *IEEE Trans. Smart Grid*, vol. 14, no. 1, pp. 4–16, Jan. 2023.
[30] Y. Jiang and Z.-P. Jiang, "Robust adaptive dynamic programming and feedback stabilization of nonlinear systems," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 25, no. 5, pp. 882–893, May 2014.
[31] 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.