Reinforcement Learning-Based Decentralized Control for Networked Interconnected Systems With Communication and Control Constraints

Jinliang Liu, *Member, IEEE*, Nan Zhang, Lijuan Zha, Xiangpeng Xie, *Senior Member, IEEE*, and Engang Tian

***Abstract*— This paper investigates the optimal decentralized control issue for a class of networked interconnected systems (NISs) under communication and control constraints. First of all, the decentralized optimal control problem for NISs is transformed into the optimal control issue for nominal systems. And then the event-triggered Hamilton-Jacobi-Isaac (HJI) equations are derived with optimal control theory based on the constructed cost function and adopted adaptive event-triggered mechanism (AETM). Furthermore, a Reinforcement Learning (RL) based event-triggered optimal control algorithm with actor-critic networks is proposed to obtain numerical solutions of the HJI equations and the Uniformly Ultimately Bounded (UUB) stability of NISs is proved to be guaranteed with Lyapunov stability theorem. Eventually, a simulation experiment is conducted to verify the effectiveness of proposed RL-based optimal decentralized algorithm with AETM.**

***Note to Practitioners*—Actually, the limitation of communication resource is one of the key factors leading to performance degradation of NISs. It is crucial to cope with this issue in the control of NISs. Employing proper communication protocol is an effective way to diminish the adverse effects of undesirable phenomena, such as data loss and out-of-order. In addition, given the scale is gradually expanded and the structure is gradually complex for NISs, the operation cost is a problem that cannot be ignored in the control process of NISs. Thus, it is significant and challenging to minimize the operation cost on the basis of ensuring system stability. In view of the communication constraints and optimization problem, the HJI equations based on AETM are derived to model the decentralized optimal control problem for NISs. Moreover, a RL-based iterative algorithm is proposed to derive the numerical solutions to HJI equations and further effectively stabilize the targeted control constrained NISs in the sense of UUB.**

***Index Terms*— Adaptive event-triggered mechanism, actor-critic network, control constraints, networked interconnected systems, reinforcement learning, uniformly ultimately bounded.**

Manuscript received 15 June 2023; accepted 27 July 2023. Date of publication 22 November 2023; date of current version 8 August 2024. This article was recommended for publication by Associate Editor M. Franceschelli and Editor C. Seatzu upon evaluation of the reviewers' comments. This work was supported in part by the National Natural Science Foundation of China under Grant 62373252, Grant 62273174, and Grant 61973152, in part by the Natural Science Foundation of Jiangsu Province of China under Grant BK20211290 and Grant BK20230063, and in part by the Qing Lan Project. *(Corresponding author: Lijuan Zha.)*

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).

Nan Zhang and Lijuan Zha are with the College of Information Engineering, Nanjing University of Finance and Economics, Nanjing, Jiangsu 210023, China (e-mail: zhangnan_nufe@163.com; zhalijuan@vip.163.com).

Xiangpeng Xie is with the Institute of Advanced Technology, Nanjing University of Posts and Telecommunications, Nanjing, Jiangsu 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).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TASE.2023.3300917.

Digital Object Identifier 10.1109/TASE.2023.3300917

I. Introduction

As IS well known, networked control systems (NCSs) have been one of the most focused researching issues in the robust control field since the special structure that system components are connected via communication network [1], [2], [3], [4]. With the widespread applications in diverse environments, networked interconnected systems (NISs), in which multiple subsystems are geographically distributed and physically interconnected, gradually attract more and more attentions [5]. To cost-effectively guarantee the stable operation of NISs, many researches on studying decentralized control have been reported [6], [7], [8], [9]. For example, the authors in [6] investigated decentralized control issue for a class of nonlinear large scale interconnected systems via Echo State Network (ESN). Actually, the practical control signal is generally constrained by a certain threshold, which further complicates the decentralized control problem for NISs [10], [11], [12], [13]. In view of this, some investigations have been conducted to settle the control constraints problem for NISs. For instance, a novel distributed event-triggered H_∞ control algorithm was proposed in [11] to achieve the expected performance for NISs with control constraints.

It is noteworthy that limited communication resources and bandwidth cannot meet the large scale data transmission in complicated NCSs, which is still a challenging issue nowadays. The undesirable phenomena and the system performance degradation are unavoidable if the constrained communication resources and bandwidth issue are not dealt with well. To overcome this issue, many researchers devote to design proper data transport scheme, among which event-triggered mechanisms have been widely applied [14], [15], [16], [17], [18], [19]. Under event-triggered mechanisms, the data can be released into the communication network only while the fluctuation of system states exceeds a pre-defined threshold. However, it is hard to determine the event-triggering threshold properly. Thus many researchers propose adaptive event-triggered mechanism (AETM) with dynamic threshold which can be adjusted adaptively according to the real-time state fluctuation of the system [20], [21], [22], [23], [24]. Based on AETM, many relative investigations have been done. For instance, the authors in [22] designed an adaptive observer-based event-triggered fault-tolerant controller to address the synchronization problem for multi-agents systems. Besides, the output feedback control problem was settled in [24] under AETM for nonlinear systems with unknown polynomial-of-output growth rate. In this paper, an AETM will be adopted to arrange the data transmission and save the scare communication resources in NISs.

During the progress of designing proper control strategies for NCSs, the whole control cost is a nonnegligible problem as revealed in recent studies. Therefore, the optimal control with the aim to minimize the whole control cost attracts many scholars' attentions [25], [26], [27], [28]. A general method to tackle the optimal control issue is to transform it into solving Hamilton-Jacobi-Isaac (HJI) equation for target nonlinear NCSs. Nevertheless, there is a bottleneck that it is hard to obtain the closed form solution to the HJI equation directly. For this, adaptive dynamic programming (ADP) [29], [30] and reinforcement learning (RL) algorithms [31], [32], [33], [34], [35], [36] have been proposed to derive the numerical solution of the HJI equation. For example, the authors in [29] developed a novel iterative ADP algorithm to solve the optimal impulsive control problem for discrete-time nonlinear systems. For RL algorithms, the basic idea is to design proper neural network and tuning law to obtain a solution to the HJI equation. Up to now, extensive research results have been achieved by using RL algorithms. For example, in [34], the actor-critic network was utilized to solve the optimal fault tolerant control issue for uncertain nonlinear MIMO system. Nevertheless, to our best knowledge, the RL-based optimal decentralized control for constrained input NISs under AETM has not been explored.

Inspired by the analysis above, this paper mainly investigates the optimal decentralized control issue for a class of NISs under AETM and control constraints. The main contributions of this paper are summarized as following:

- A RL-based optimal decentralized control algorithm with AETM is proposed to stabilize the NISs. Compared with [37], the restrictive condition of the nonlinear systems with zero equilibrium is removed in this paper, which makes the proposed RL-based optimal decentralized control algorithm applicable for nonlinear systems that have non-zero equilibrium points;
- Different from some optimal control algorithms by applying the static event-triggered scheme [38], the AETM in this article is adopted to settle communication constraints for NISs and the triggering threshold can be adaptively adjusted according to real-time state operation of systems.
- The actor and critic networks are utilized to obtain numerical solutions of the HJI equations with the help of gradient descent method and concurrent learning technique.

The remaining paper is organized as follows. In Section II, based on the formulated NISs, the optimal decentralized control problem is transformed into optimal control problems of nominal subsystems. Then the event-based HJI equations are derived with adopted AETM. In Section III, a RL-based event-triggered iterative algorithm is designed to derive the numerical solutions of the HJI equations and the effectiveness of which is proved. The performance of proposed RL-based algorithm is testified by a experiment in Section IV and Section V summaries this paper.

II. Problem Statement

In this section, the considered NISs with control constraints is described firstly. Then, based on the established nominal systems and the designed AETM, the studied distributed optimal control problem is transformed into solving a set of event-triggered HJI equations.

A. System Description

We consider a continuous time NISs consisting of N subsystems and the ith subsystem S_i $(i = 1, 2, \ldots, N)$ can be depicted as following:

$$\dot{x}_i(t) = f_i(x_i(t)) + g_i(x_i(t))u_i(t) + h_i(x_i(t))\omega_i(x(t)), \quad (1)$$

in which $x_i(t) \in \mathcal{R}^n$ and $u_i(t) \in \mathcal{R}^m$ denote the system state and the control input of S_i, respectively; $\omega_i(x(t)) \in \mathcal{R}^q$ is the nonlinear interconnected term with other subsystems. Besides, $f_i(x_i(t)) : \mathcal{R}^n \to \mathcal{R}^n$, $g_i(x_i(t)) : \mathcal{R}^n \to \mathcal{R}^{n\times m}$ and $h_i(x_i(t)) : \mathcal{R}^n \to \mathcal{R}^{n\times q}$ are known nonlinear smooth mappings, respectively.

Remark 1: As is illustrated in [39], in order to simulate the actual situation, this paper mainly investigates the unmatched interconnection with $g_i(x_i(t)) \neq h_i(x_i(t))$.

Assumption 1: [35]: For each S_i, the nonlinear interconnected coupled term $\omega_i(x(t))$ is set to be bounded as:

$$||\omega_i(x(t))|| \le \sum_{j=1}^{N} a_{ij} P_{ij}(x_j(t)), \quad (2)$$

in which $a_{ij} > 0$ and positive function $P_{ij}(x_j(t)) \in \mathcal{R}$ represent the coupled interconnections.

Denote $P_i(x_i(t)) = max\{P_{1i}(x_i(t)), \ldots, P_{Ni}(x_i(t))\}$, then condition (2) can be represented as:

$$||\omega_i(x(t))|| \le \sum_{j=1}^{N} b_{ij} P_j(x_j(t)), \quad (3)$$

where $b_{ij} \ge a_{ij} P_{ij}(x_j(t))/P_j(x_j(t))$.

Given that the control constraint is considered in this study, the following non-quadratic function is introduced to depict the constrained input [35]:

$$\Pi_i(u_i(t)) = 2\lambda_i \int_0^{u_i(t)} \Gamma^{-T}(v/\lambda_i) R_i dv, \quad (4)$$

where $u_i(t)$ is the feedback control of S_i with $|u_i(t)| \le \lambda_i$; $R_i = diag\{r_{i1}, \ldots, r_{im}\}$ is the positive definite matrix and $\Gamma(\cdot) = tanh(\cdot)$ is the hyperbolic tangent function. Then according to the integral calculation result and using the similar methods in [40] and [41], the Eq. (4) can be transformed into:

$$\Pi_i(u_i(t)) = 2\lambda_i u_i^T(t) R_i \Gamma^{-1}(u_i(t)/\lambda_i) + \lambda_i^2 \tilde{R}_i ln(\mathbf{1} - (u_i(t)/\lambda_i)^2), \quad (5)$$

where $\mathbf{1} = [1, \ldots, 1]^T \in \mathcal{R}^{m\times 1}$ and $\tilde{R}_i = [r_{i1}, \ldots, r_{im}] \in \mathcal{R}^{1\times m}$.

Definition 1 [42]: The system state $x(t)$ is said to be stable in the sense of Uniformly Ultimately Bounded (UUB), if there exists a compact set $\Omega \subset \mathbb{R}^n$, a constant θ and a time instant $t_f = \mathbb{T}(\theta, z_0)$ such that $||x(t)|| \le \theta$ for all $t \ge t_0 + t_f$ with initial value $x(t_0) = x_0 \in \Omega$, in which x_0 denotes the initial state of the system.

Based on the formulated system, the main objective is to design an optimal state-feedback decentralized control strategy to guarantee that the NISs is stable in the sense of UUB. Nevertheless, the aim is hard to be achieved due to the existence of interconnected term. In view of this, by referring to [43] and [44], we then transform the decentralized stabilization problem into solving N optimal control subproblems over nominal subsystems corresponding to the NISs (1).

B. Problem Formulation

Given that the N subsystems in the NISs are homogeneous, without loss of generality, the following introduction in the subsection is focused on S_i $(i = 1, 2, \ldots, N)$. The continuous time nonlinear nominal subsystem S_i' corresponding to S_i can be formulated as:

$$\dot{x}_i(t) = f_i(x_i(t)) + g_i(x_i(t))u_i(t), \tag{6}$$

then the infinite-horizon cost function for S_i' under control input constraint is defined as:

$$V_i(x_i(t)) = \int_t^{\infty} [e^{-\alpha_i(\tau - t)}(x_i(t)^T Q_i x_i(t) + \Pi_i(u_i(t)) + \eta_i P_i^2(x_i(t)))]d\tau, \tag{7}$$

where α_i is the given discount factor which is designed to be bounded by a constant a, i.e., $\alpha_i < a$, and $\eta_i > 0$ is the adjusting parameter; $Q_i \in \mathcal{R}^{n\times n}$ is the known positive-definite matrix.

Based on the above cost function, the optimal control cost $V_i^*(x_i(t))$ is described as:

$$V_i^*(x_i(t)) = \min_{u_i(t)\in\Omega(\Phi_i)} V_i(x_i(t)), \tag{8}$$

in which $\Omega(\Phi_i)$ denotes the set of all admissible control policies designed on Φ_i.

According to optimal control theory, $V_i^*(x_i(t))$ can be derived by solving the following HJI equations:

$$\min_{u_i(t)\in\Omega(\Phi_i)} H_i(x_i(t), V_i^*(x_i(t)), u_i(t)) = 0, \tag{9}$$

in which $H_i(x_i(t), V_i^*(x_i(t)), u_i(t))$ denotes the Hamilton function for S_i' and can be specifically presented as:

$$\begin{aligned} &H_i(x_i(t), V_i^*(x_i(t)), u_i(t)) \\ &= (\nabla V_i^*(x_i(t)))^T (f_i(x_i(t)) + g_i(x_i(t))u_i(t)) + \eta_i P_i^2(x_i(t)) \\ &\quad + x_i(t)^T Q_i x_i(t) + \Pi_i(u_i(t)) - \alpha_i V_i^*(x_i(t)), \end{aligned} \tag{10}$$

where $\nabla V_i^*(x_i(t))$ denotes the partial differential of $V_i^*(x_i(t))$ with respect to $x_i(t)$. Then according to stationarity condition, i.e., $\partial H_i(x_i(t), V_i^*(x_i(t)), u_i(t))/\partial u_i(t) = 0$, the optimal control input $u_i^*(x_i(t))$ can be computed as:

$$u_i^*(x_i(t)) = -\lambda_i \Gamma(\frac{1}{2\lambda_i} R_i^{-1} g_i^T(x_i(t))\nabla V_i^*(x_i(t))). \tag{11}$$

As is mentioned before, an AETM is adopted to effectively schedule data transmission in the communication network, under which the sampled data in S_i' can be released only when the following condition holds [45]:

$$z_{i,k}^T \Psi_i z_{i,k} > \zeta_i(t_k^i h + l_i h) x_i^T(t_k^i h + l_i h)\Psi_i x_i(t_k^i h + l_i h), \tag{12}$$

where $z_{i,k} = x_i(t_k^i h + l_i h) - x_i(t_k^i h)$ with $l_i = \{1, 2, 3, \ldots\}$. h is the sampling period. $x_i(t_k^i h)$ and $x_i(t_k^i h + l_i h)$ denote the transmitted data at the last event-triggering instant $t_k^i h$ and the current sampling instant $t_k^i h + l_i h$, respectively. Moreover, the adaptive parameter $\zeta_i(t_k^i h + l_i h)$ is designed as:

$$\zeta_i(t_k^i h + l_i h) = \underline{\zeta_i} + (\bar{\zeta}_i - \underline{\zeta_i})e^{-\varepsilon_i(z_{i,k}^T \Psi_i z_{i,k})}, \tag{13}$$

in which $\varepsilon_i > 0$ is adopted to adjust the sensitivity of the adaptive threshold $\zeta_i(t_k^i h + l_i h)$ to $z_{i,k}$; $\bar{\zeta}_i$ and $\underline{\zeta_i}$ are pre-defined upper and lower bounds of $\zeta_i(t_k^i h + l_i h)$, i.e., $\underline{\zeta_i} \le \zeta_i(t_k^i h + l_i h) \le \bar{\zeta}_i$.

Based on the above AETM Eq. (12), the next event-triggering instant $t_{k+1}^i h$ can be rewritten as:

$$t_{k+1}^i h = t_k^i h + \min_{l_i \ge 1}\{l_i h \mid l_i \text{ satisfies condition } Eq. (12)\}. \tag{14}$$

Remark 2: It is noted that the AETM (12) will reduce to the traditional static ETM when $\underline{\zeta_i} = \bar{\zeta}_i$. Besides, the adaptive triggering law $\zeta_i(t_k^i h + l_i h)$ is designed to be state-dependent, i.e., the large state fluctuation $z_{i,k}$ results in the small $\zeta_i(t_k^i h + l_i h)$, then the adaptive event-triggering condition Eq. (12) is more easily to be satisfied, and vice versa. Moreover, the parameter ε_i denotes the gradient of exponential function which can adjust the sensitivity of $\zeta_i(t_k^i h + l_i h)$ to the state fluctuation $z_{i,k}$.

Under the AETM, the control input will be updated at each event-triggering instant. Then the optimal event-triggered control law can be represented as:

$$\begin{aligned} u_i^*(x_i(t)) &= u_i^*(x_i(t_k^i h)) \\ &= -\lambda_i \Gamma(\frac{1}{2\lambda_i} R_i^{-1} g_i^T(x_i(t_k^i h))\nabla V_i^*(x_i(t_k^i h))). \end{aligned} \tag{15}$$

Consequently, the event-triggered HJI equations can be given as:

$$\begin{aligned} &H_i(x_i(t), V_i^*(x_i(t)), u_i^*(x_i(t_k^i h))) \\ &= (\nabla V_i^*(x_i(t)))^T (f_i(x_i(t)) + g_i(x_i(t))u_i^*(x_i(t_k^i h))) \\ &\quad + \eta_i P_i^2(x_i(t)) + x_i(t)^T Q_i x_i(t) \\ &\quad + \Pi_i(u_i^*(x_i(t_k^i h))) - \alpha_i V_i^*(x_i(t)) \\ &= 0. \end{aligned} \tag{16}$$

III. Main Results

In this section, the stability of the NISs under the envisioned optimal control strategy is analyzed firstly. Then, a RL-based algorithm is designed to obtain an appropriate solution to the studied control problem. The effectiveness of the algorithm is finally validated based on Lyapunov theorem.[1]

[1]For convenience of the further analysis, denoting $x_i(t)$, $x_i(t_k^i h)$, $V_i^*(x_i(t))$, $\nabla V_i^*(x_i(t))$ as x_i, $x_{i,k}$, V_i^*, $V_{x_i}^*$, respectively.

A. Stability Analysis

Before giving the specific analysis, we would like to introduce the following assumptions.

Assumption 2 [35]: $||g_i(x_i)|| < g_{im}$ and $||h_i(x_i)|| < h_{im}$, where g_{im} and h_{im} are positive constants.

Assumption 3 [46]: V_i^* and $V_{x_i}^*$ are bounded by positive constants V_{im} and V_{xm}, i.e., $||V_i^*|| < V_{im}$ and $||V_{x_i}^*|| < V_{xm}$.

Assumption 4 [47]: $\mathcal{D}_i^*$ is assumed to be Lipschitz continuous on Ω_i and satisfy:

$$||\mathcal{D}_i^*(x_i) - \mathcal{D}_i^*(x_{i,k})|| \le \mathcal{L}_i||x_i - x_{i,k}|| \triangleq \mathcal{L}_i||z_{i,k}||, \tag{17}$$

where $\mathcal{D}_i^*(x_{i,k}) = \frac{1}{2\lambda_i}R_i^{-1}g_i^T(x_{i,k})V_{x_{i,k}}^*$ and $\mathcal{L}_i$ is a positive real constant.

Remark 3: In Assumption 2, $f_i(x_i)$ and $g_i(x_i)$ are assumed to have the upper bounds, which is motivated by [35]. This assumption makes the proposed algorithm more general than that used in the existing work.

Remark 4: Assumption 3 implies that V_i^* and $V_{x_i}^*$ are bounded. Actually, it should be noted that the optimal value function V_i^* is continuously differentiable on its admissible set according to literature [46], which means that both V_i^* and $V_{x_i}^*$ are bounded on their admissible sets.

Remark 5: The nonlinear term $\mathcal{D}_i^*$ is assumed be Lipschitz continuous in Assumption 4, which is usually used in literatures such as [47], i.e., lots of nonlinear models of practical systems can be described as Lipschitz. For example, the sinusoidal functions are globally Lipschitz and often appear in robotic systems [35].

Then, based on the analysis above and Assumptions 1-4, the UUB stability of considered NISs under proposed optimal control strategies is proved by the following theorem.

Theorem 1: Given the optimal cost functions V_i^* ($i = 1, 2, \dots, N$) designed with Eq. (8), the considered NISs can be stabilized in the sense of UUB under optimal control laws $u_i^*(x_{i,k})$ depicted by Eq. (15) and Assumptions 1-4, if there exist positive constants η_i^* such that $\eta_i > \eta_i^*$, and the following conditions hold:

$$\lambda_{\min}(Q_i) - 2\lambda_i^2\mathcal{L}_i^2||R_i||\bar{\zeta}_i\frac{\lambda_{\max}(\Psi_i)}{\lambda_{\min}(\Psi_i)} \ge 0. \tag{18}$$

Proof: The following Lyapunov function is constructed:

$$L(x) = \sum_{i=1}^{N} V_i^*. \tag{19}$$

Then differentiating $L(x)$ with respect to time t under system equations: $\dot{x}_i = f_i(x_i) + g_i(x_i)u_i^*(x_{i,k}) + h_i(x_i)\omega_i(x)$, it can be gotten that:

$$\dot{L}(x) = \sum_{i=1}^{N}\{(V_{x_i}^*)^T(f_i(x_i) + g_i(x_i)u_i^*(x_{i,k}) + h_i(x_i)\omega_i(x))\}. \tag{20}$$

According to Eq. (10), it is obtained that:

$$\begin{aligned}(V_{x_i}^*)^T f_i(x_i) &= \alpha_i V_i^* - x_i^T Q_i x_i - \eta_i P_i^2(x_i) \\ &\quad - (V_{x_i}^*)^T g_i(x_i)u_i^*(x_i) - \Pi_i(u_i^*(x_i)).\end{aligned} \tag{21}$$

Based on Eq. (5), $\Pi_i(u_i^*(x_i))$ can be rewritten as:

$$\begin{aligned}\Pi_i(u_i^*(x_i)) &= 2\lambda_i(u_i^*(x_i))^T R_i\Gamma^{-1}(u_i^*(x_i)/\lambda_i) \\ &\quad + \lambda_i^2\tilde{R}_i ln(\mathbf{1} - (u_i^*(x_i)/\lambda_i)^2).\end{aligned} \tag{22}$$

Moreover, according to the event-triggered optimal control law (15) and referring to the similar method in [35], the Eq. (20) can be rewritten as:

$$\begin{aligned}\dot{L}(x) &= \sum_{i=1}^{N}\{\alpha_i V_i^* - x_i^T Q_i x_i - \eta_i P_i^2(x_i) \\ &\quad + \int_{u_i^*(x_i)}^{u_i^*(x_{i,k})} 2\lambda_i[\Gamma^{-1}(\upsilon/\lambda_i) + \mathcal{D}_i^*(x_i)]^T R_i d\upsilon \\ &\quad - \Pi_i(u_i^*(x_{i,k})) + (V_{x_i}^*)^T h_i(x_i)\omega_i(x)\}.\end{aligned} \tag{23}$$

Defining $\upsilon = -\lambda_i\Gamma(\varpi)$, it can be obtained that:

$$\begin{aligned}&\int_{u_i^*(x_i)}^{u_i^*(x_{i,k})} 2\lambda_i[\Gamma^{-1}(\upsilon/\lambda_i) + \mathcal{D}_i^*(x_i)]^T R_i d\upsilon \\ &\le \int_{\mathcal{D}_i^*(x_i)}^{\mathcal{D}_i^*(x_{i,k})} 4\lambda_i^2(\upsilon - \mathcal{D}_i^*(x_i))^T R_i d\upsilon \\ &= 2\lambda_i^2(\mathcal{D}_i^*(x_{i,k}) - \mathcal{D}_i^*(x_i))^T R_i(\mathcal{D}_i^*(x_{i,k}) - \mathcal{D}_i^*(x_i)) \\ &\le 2\lambda_i^2\mathcal{L}_i^2||R_i||||z_{i,k}||^2.\end{aligned} \tag{24}$$

Then with the help of Young's inequality, Eq. (23) can be rewritten as:

$$\begin{aligned}&\dot{L}(x) \\ &\le \sum_{i=1}^{N}\{\alpha_i V_i^* - x_i^T Q_i x_i + 2\lambda_i^2\mathcal{L}_i^2||R_i||||z_{i,k}||^2 \\ &\quad - \Pi_i(u_i^*(x_{i,k})) + \frac{1}{2}V_{xm}^2 h_{im}^2 - (\eta_i P_i^2(x_i) - \frac{1}{2}\omega_i^T(x)\omega_i(x))\}.\end{aligned} \tag{25}$$

In order to handle the interconnected term in Eq. (25), we define the following items:

$$\begin{cases}\tilde{\eta} = diag\{\eta_1, \eta_2, \cdots, \eta_N\} \\ \tilde{\mathbf{1}} = diag\{1_1, 1_2, \cdots, 1_N\} \\ y(x) = [-P_1(x_1), -P_2(x_2), \cdots, -P_N(x_N), \\ \quad 1_1, 1_2, \cdots, 1_N]^T\end{cases} \tag{26}$$

Then Eq. (25) yields that:

$$\begin{aligned}\dot{L}(x) &\le \sum_{i=1}^{N}\{\alpha_i V_i^* - x_i^T Q_i x_i + 2\lambda_i^2\mathcal{L}_i^2||R_i||||z_{i,k}||^2 \\ &\quad - \Pi_i(u_i^*(x_{i,k})) + \frac{1}{2}V_{xm}^2 h_{im}^2\} - y^T(x)\tilde{S}y(x) - N,\end{aligned} \tag{27}$$

where

$$\tilde{S} = \begin{bmatrix}\tilde{\eta} & \frac{1}{4}BB^T \\ \frac{1}{4}B^T B & \tilde{\mathbf{1}}\end{bmatrix} \text{ and } B = \begin{bmatrix}b_{11} & b_{12} & \cdots & b_{1N} \\ \vdots & \vdots & \ddots & \vdots \\ b_{N1} & b_{N2} & \cdots & b_{NN}\end{bmatrix}.$$

According to the definition of $\tilde{S}$, it can be found that $\tilde{S}$ is positive definite if the values of η_i ($i = 1, 2, \dots, N$) are properly selected. In other words, if there exist N positive constants η_i^* such that $\eta_i > \eta_i^*$, then we have $y^T(x)\tilde{S}y(x) > 0$.

In addition, based on the AETM (12), and the adaptive parameters $\zeta_i(t_k^i h + l_i h)$, it can be derived that ($t \in [t_k^i h, t_{k+1}^i h)$):

$$\lambda_{\min}(\Psi_i)||z_{i,k}||^2 \leq \bar{\zeta}_i \lambda_{\max}(\Psi_i)||x_i(t_k^i h + l_i h)||^2. \tag{28}$$

Based on the analysis above, it yields that:

$$\begin{aligned}\dot{L}(x) \leq & \sum_{i=1}^{N}\{\alpha_i V_i^* - x_i^T Q_i x_i - \Pi_i(u_i^*(x_{i,k})) + \frac{1}{2}V_{xm}^2 h_{im}^2 \\ & + 2\lambda_i^2 \mathcal{L}_i^2 ||R_i||\bar{\zeta}_i \frac{\lambda_{\max}(\Psi_i)}{\lambda_{\min}(\Psi_i)}||x_i(t_k^i h + l_i h)||^2\} \\ \leq & \sum_{i=1}^{N}\{aV_{im} + \frac{1}{2}V_{xm}^2 h_{im}^2 \\ & + (2\lambda_i^2 \mathcal{L}_i^2 ||R_i||\bar{\zeta}_i \frac{\lambda_{\max}(\Psi_i)}{\lambda_{\min}(\Psi_i)} - \lambda_{\min}(Q_i))||x_i||^2\}. \end{aligned} \tag{29}$$

It can be found that $\dot{L}(x) < 0$ when the conditions (18) hold and $x_i \notin \mathcal{Q}_{x_i}$, where $\mathcal{Q}_{x_i}$ is defined as:

$$\mathcal{Q}_{x_i} = \left\{ x_i : ||x_i|| \leq \sqrt{\frac{aV_{im} + \frac{1}{2}V_{xm}^2 h_{im}^2}{\lambda_{\min}(Q_i) - 2\lambda_i^2 \mathcal{L}_i^2 ||R_i||\bar{\zeta}_i \frac{\lambda_{\max}(\Psi_i)}{\lambda_{\min}(\Psi_i)}}} \right\}. \tag{30}$$

Thus, it implies that with the holding of conditions (18), there exists a time instant t_f such that x_i can converge to a certain bound when $t > t_0 + t_f$. So the NISs (1) is stable in the sense of UUB according to Definition 1. ■

Remark 6: In terms of time complexity of the algorithm proposed in this paper, considering that the RL-based algorithm uses two NNs, namely the actor NN and the critic NN, the time complexity is $O(n^2)$ with n denoting the number of iterations. Moreover, given that the algorithm computation depends on dynamics information of the systems and the hyperbolic tangent function, then the overall time complexity of the algorithm proposed in this paper is $O(mn^3)$, with m denoting the dimension of system vector.

B. Actor-Critic RL Algorithm

Given that it is hard to obtain the close-form solutions of the event-triggered HJI equations (16), then in this subsection, a RL algorithm is proposed to derive the approximated solutions to the HJI equations. In the algorithm, a critic and an actor networks are utilized to model each optimal value function V_i^* ($i = 1, 2, \ldots, N$) and optimal constrained control input $u_i^*(x_{i,k})$, respectively. The following description and analysis for the algorithm is focused on $t \in [t_k^i h, t_{k+1}^i h)$ unless noted otherwise.

1) Critic Network Design: The optimal cost function V_i^* can be represented by the following critic network:

$$V_i^* = W_{ci}^T \phi_{ci}(x_i) + \varepsilon_{ci}(x_i), \tag{31}$$

in which $W_{ci} \in \mathcal{R}^{n_{ci}}$ is the optimal weight vector to be designed (n_{ci} denotes the number of the hidden neurons); $\phi_{ci}(x_i) = col\{\phi_{ci}(x_{ip})\}$, ($p = 1, 2, \ldots n_{ci}$) is the pre-defined activation function, and $\phi_{ci}(x_{ip})$ ($i = 1, 2, \ldots, N$) are independent with each other; $\varepsilon_{ci}(x_i)$ is the critic residual error. Then the partial derivative of $V_i^*(x_i)$ with respect to x_i can be computed as:

$$V_{x_i}^* = \nabla\phi_{ci}^T(x_i)W_{ci} + \nabla\varepsilon_{ci}(x_i). \tag{32}$$

In practice, considering that W_{ci} is hardly available, we then use an estimated value $\hat{W}_{ci}$ to replace W_{ci}, and thereby get the approximation of V_i^* as:

$$\hat{V}_i = \hat{W}_{ci}^T \phi_{ci}(x_i). \tag{33}$$

Consequently, the partial derivative of $\hat{V}_i$ with respect to x_i can be calculated as:

$$\hat{V}_{x_i} = \nabla\phi_{ci}^T(x_i)\hat{W}_{ci}. \tag{34}$$

On the basis of Eqs. (31)-(34), the approximated Hamilton function $\hat{H}_i(x_i, \hat{V}_i(x_i), \hat{u}_i(x_{i,k}))$ can be written as:

$$\begin{aligned}&\hat{H}_i(x_i, \hat{V}_i(x_i), \hat{u}_i(x_{i,k})) \\ &= \hat{W}_{ci}^T \nabla\phi_{ci}(x_i)(f_i(x_i) + g_i(x_i)\hat{u}_i(x_{i,k})) \\ &\quad + \eta_i P_i^2(x_i) + x_i^T Q_i x_i + \Pi_i(\hat{u}_i(x_{i,k})) \\ &\quad - \alpha_i \hat{W}_{ci}^T \phi_{ci}(x_i). \end{aligned} \tag{35}$$

Denoting the error between $H_i(x_i, V_i^*(x_i), u_i^*(x_{i,k}))$ and $\hat{H}_i(x_i, \hat{V}_i(x_i), \hat{u}_i(x_{i,k}))$ as φ_{ci}, then it can be obtained that:

$$\begin{aligned}\varphi_{ci} &= \hat{H}_i(x_i, \hat{V}_i(x_i), \hat{u}_i(x_{i,k})) - H_i(x_i, V_i^*(x_i), u_i^*(x_{i,k})) \\ &= \hat{W}_{ci}^T \vartheta_i + \psi_{ci}, \end{aligned} \tag{36}$$

where

$$\begin{cases} \vartheta_i = \nabla\phi_{ci}(x_i)(f_i(x_i) + g_i(x_i)\hat{u}_i(x_{i,k})) - \alpha_i\phi_{ci}(x_i), \\ \psi_{ci} = \eta_i P_i^2(x_i) + x_i^T Q_i x_i + \Pi_i(\hat{u}_i(x_{i,k})). \end{cases}$$

The main objective of the paper is to make $\hat{u}_i(x_{i,k}) \to u_i^*(x_{i,k})$, i.e., to make $\varphi_{ci} \to 0$. Towards this end, one general method is to minimize the target function E_{c_i}. In order to achieve a high efficiency in utilizing the historical state data [35], we then design the following target function:

$$\begin{aligned}E_{c_i} &= E_{ci} + \sum_{s=1}^{N_i(t)} e^{t_s^i h - t} E_{ci,s} \\ &= \frac{1}{2}\varphi_{ci}^T \varphi_{ci} + \frac{1}{2}\sum_{s=1}^{N_i(t)} e^{t_s^i h - t}\varphi_{ci,s}^T \varphi_{ci,s}, , \end{aligned} \tag{37}$$

where $N_i(t)$ is the number of event-triggering instants for S_i' in time interval $[0, t]$, and $\varphi_{ci,s}$ denotes the historical data at time $t_s^i h$:

$$\varphi_{ci,s} = \hat{W}_{ci}^T \vartheta_{i,s} + \psi_{ci,s}, \tag{38}$$

where

$$\begin{cases} \vartheta_{i,s} = \nabla\phi_{ci}(x_{i,s})(f_i(x_{i,s}) + g_i(x_{i,s})\hat{u}_i(x_{i,k})) - \alpha_i\phi_{ci}(x_{i,s}), \\ \psi_{ci,s} = \eta_i P_i^2(x_{i,s}) + x_{i,s}^T Q_i x_{i,s} + \Pi_i(\hat{u}_i(x_{i,k})). \end{cases}$$

Then the tuning law of the approximated weight vector $\hat{W}_{ci}$ is designed as:

$$\dot{\hat{W}}_{ci} = -\frac{\iota_{ci}}{(1 + \vartheta_i^T \vartheta_i)^2}\frac{\partial E_{ci}}{\partial \hat{W}_{ci}} - \sum_{s=1}^{N_i(t)} \frac{\iota_{ci}}{(1 + \vartheta_{i,s}^T \vartheta_{i,s})^2}\frac{\partial E_{ci,s}}{\partial \hat{W}_{ci}}$$

$$= -\frac{\iota_{ci}\vartheta_i}{(1+\vartheta_i^T\vartheta_i)^2}\varphi_{ci} - \sum_{s=1}^{N_i(t)} \frac{\iota_{ci}\vartheta_{i,s}e^{t_s^i h - t}}{(1+\vartheta_{i,s}^T\vartheta_{i,s})^2}\varphi_{ci,s}, \quad (39)$$

in which $\iota_{ci} \in (0,1)$ implies the learning rate which will affect the convergence of $\hat{W}_{ci}$. Defining $\tilde{W}_{ci} = W_{ci} - \hat{W}_{ci}$, then it can be gotten:

$$\dot{\tilde{W}}_{ci} = -\iota_{ci}\Big(\frac{\vartheta_i\vartheta_i^T}{(1+\vartheta_i^T\vartheta_i)^2} + \sum_{s=1}^{N_i(t)} \frac{\vartheta_{i,s}\vartheta_{i,s}^T e^{t_s^i h - t}}{(1+\vartheta_{i,s}^T\vartheta_{i,s})^2}\Big)\tilde{W}_{ci} + \frac{\iota_{ci}\vartheta_i}{(1+\vartheta_i^T\vartheta_i)^2}\varrho_{ci} + \sum_{s=1}^{N_i(t)} \frac{\iota_{ci}\vartheta_{i,s}e^{t_s^i h - t}}{(1+\vartheta_{i,s}^T\vartheta_{i,s})^2}\varrho_{ci,s}, \quad (40)$$

where ϱ_{ci} and $\varrho_{ci,s}$ are the residual errors which can be represented as:

$$\begin{cases} \varrho_{ci} = -\nabla\varepsilon_{ci}^T(x_i)(f_i(x_i) + g_i(x_i)\hat{u}_i(x_{i,k})) + \alpha_i\varepsilon_{ci}(x_i), \\ \varrho_{ci,s} = -\nabla\varepsilon_{ci}^T(x_{i,s})(f_i(x_{i,s}) + g_i(x_{i,s})\hat{u}_i(x_{i,k})) \\ \qquad +\alpha_i\varepsilon_{ci}(x_{i,s}). \end{cases}$$

Remark 7: Motivated by the work of [35], the concurrent learning technique and gradient decent method are adopted in designing the updating policy of $\hat{W}_{ci}$. But different from the referred literature, the decry term $e^{t_s^i h - t}$ is introduced in Eq. (37), then the newer historical system states will play more important role in the designed objective function E_{ci}.

2) Actor Network Design: In this paper, an actor network is utilized to depict each optimal event-triggered control input $u_i^*(x_{i,k})$, to be specific:

$$u_i^*(x_{i,k}) = W_{ui}^T\phi_{ui}(x_{i,k}) + \varepsilon_{ui}(x_{i,k}), \quad (41)$$

in which $W_{ui} \in \mathcal{R}^{n_{ui}}$ and $\phi_{ui}(x_i) = col\{\phi_{ui}(x_{ip})\}, (p = 1, 2, \ldots n_{ui})$ are the optimal weight vector and basis function, respectively; $\varepsilon_{ui}(x_i)$ is the actor residual error. Similar to the analysis above, $u_i^*(x_{i,k})$ can be approximated as:

$$\hat{u}_i(x_{i,k}) = \hat{W}_{ui}^T\phi_{ui}(x_{i,k}), \quad (42)$$

where $\hat{W}_{ui} \in \mathcal{R}^{n_{ui}}$ denotes the estimated value of W_{ui}.

Inspired by [47], the actor critic error φ_{ui} is then designed as follows:

$$\varphi_{ui} = \hat{u}(x_{i,k}) - u_{\hat{W}_{ci}} = \hat{W}_{ui}^T\phi_{ui}(x_{i,k}) + \lambda_i\Gamma\Big(\frac{1}{2\lambda_i}R_i^{-1}g_i^T(x_{i,k})\nabla\phi_{ci}^T(x_{i,k})\hat{W}_{ci}\Big), \quad (43)$$

in which $u_{\hat{W}_{ci}}$ is the control policy (15) approximated by the evaluation of critic weight $\hat{W}_{ci}$ in Eq.(34), that is to say:

$$u_{\hat{W}_{ci}} = -\lambda_i\Gamma\Big(\frac{1}{2\lambda_i}R_i^{-1}g_i^T(x_{i,k})\hat{V}_{x_{i,k}}\Big) = -\lambda_i\Gamma\Big(\frac{1}{2\lambda_i}R_i^{-1}g_i^T(x_{i,k})\nabla\phi_{ci}^T(x_{i,k})\hat{W}_{ci}\Big), \quad (44)$$

then Eq. (43) can be obtained.

To achieve $\varphi_{ui} \to 0$, the actor network weight vector $\hat{W}_{ui}$ is generally selected with the aim to minimize the squared error performance $E_{ui} = \frac{1}{2}\varphi_{ui}^T\varphi_{ui}$. Moreover, considering that the control inputs are only updated at event-triggering instants with the employed AETM, the tuning law for $\hat{W}_{ui}$ is thereby designed as:

$$\begin{cases} \dot{\hat{W}}_{ui} = 0, \ t \in [t_k^i h, t_{k+1}^i h), \\ \hat{W}_{ui}^+ = \hat{W}_{ui} - \iota_{ui}\frac{\partial E_{ui}}{\partial \hat{W}_{ui}} = \hat{W}_{ui} - \iota_{ui}\phi_{ui}(x_{i,k}) \\ [\hat{W}_{ui}^T\phi_{ui}(x_{i,k}) + \lambda_i\Gamma(\frac{1}{2\lambda_i}R_i^{-1}g_i^T(x_{i,k}) \\ \nabla\phi_{ci}^T(x_{i,k})\hat{W}_{ci})]^T, \ t = t_{k+1}^i h. \end{cases} \quad (45)$$

where $\iota_{ui} \in (0,1)$ is the learning rate of the actor network.

Denoting $\tilde{W}_{ui} = W_{ui} - \hat{W}_{ui}$, then it can be obtained:

$$\begin{cases} \dot{\tilde{W}}_{ui} = 0, \ t \in [t_k^i h, t_{k+1}^i h), \\ \tilde{W}_{ui}^+ = \tilde{W}_{ui} - \iota_{ui}\phi_{ui}(x_{i,k})\varepsilon_{ui}^T(x_{i,k}) \\ \qquad -\iota_{ui}\phi_{ui}(x_{i,k})\phi_{ui}^T(x_{i,k})\tilde{W}_{ui} \\ \qquad -\lambda_i\iota_{ui}\phi_{ui}(x_{i,k})[\Gamma(\hat{\mathcal{D}}_i(x_{i,k})) - \Gamma(\mathcal{D}_i^*(x_{i,k}))], \\ \qquad t = t_{k+1}^i h. \end{cases} \quad (46)$$

The workflow of the RL-based optimal control algorithm is depicted in Fig. 1. First of all, whether the current sampled data can be released is determined by the event-triggered condition in Eq. (12). If the newly sampled data is transmitted, then update $\hat{u}_i(t)$ based on Eq. (42) with $\hat{W}_{ui}(t)$, $\hat{W}_{ci}(t)$ and the historical data collection. If not, then update $\hat{u}(t) = \hat{u}(t-h)$. Secondly, compute $\hat{W}_{ci}(t+h)$ and $x_i(t+h)$ according to Eqs. (39) and (6), respectively. Finally, it is to judge whether the time $t+h$ exceeds the given time window T. If it is, then terminate the whole progress. If not, then return to the first step.

C. The Effectiveness of The Proposed RL-Based Algorithm

In this subsection, in order to testify the effectiveness of the proposed RL-based algorithm, the state x_i of S_i' $(i = 1, 2, \ldots, N)$, the estimated error $\tilde{W}_{ci}$ and $\tilde{W}_{ui}$ are proved to be stable in the sense of UUB under the proposed algorithm at first. Besides, the following assumption is proposed for the further analysis.

Assumption 5 [47]: Supposing that there exist positive constants $\nabla\phi_{cim}$, $\nabla\varepsilon_{cim}$, ϕ_{uim}, ε_{uim} and ϱ_{im} satisfying the following inequalities $||\nabla\phi_{ci}|| < \nabla\phi_{cim}$, $||\nabla\varepsilon_{ci}|| < \nabla\varepsilon_{cim}$, $||\phi_{ui}|| < \phi_{uim}$, $||\varepsilon_{ui}|| < \varepsilon_{uim}$, $||\varrho_{ci}|| < \varrho_{im}$ and $||\varrho_{ci,k}|| < \varrho_{im}$.

Accordingly, combining with the designed Actor-Critic network and Assumptions 1-5, the UUB stability of estimated weight errors $\tilde{W}_{ci}$ and $\tilde{W}_{ui}$ under proposed RL-based iterative algorithm are proved via the following theorem.

Theorem 2: Given the estimated optimal control input $\hat{u}_i(x_{i,k})$ in Eq. (42), the tuning laws of $\hat{W}_{ci}$ in Eq. (39) and $\hat{W}_{ui}$ in Eq. (45), respectively, then the system states x_i for S_i' and the estimated weight errors $\tilde{W}_{ci}$ and $\tilde{W}_{ui}$ are considered to be stable in the sense of UUB if Assumptions 1-5 and the following conditions hold:

$$\begin{cases} \lambda_{\min}(Q_i) - 2\lambda_i^2\mathcal{L}_i^2(1+1/\gamma_i)||R_i||\frac{\bar{\zeta}_i\lambda_{\max}(\Psi_i)}{\lambda_{\min}(\Psi_i)} > 0, \\ \frac{\iota_{ci}}{2}\lambda_{min}(\mathcal{H}_i) - 2\lambda_i^2\frac{(1+\gamma_i)^2}{4\gamma_i\lambda_i^2}||R_i||^{-1}g_{im}^2\nabla\phi_{cim}^2 > 0. \end{cases} \quad (47)$$

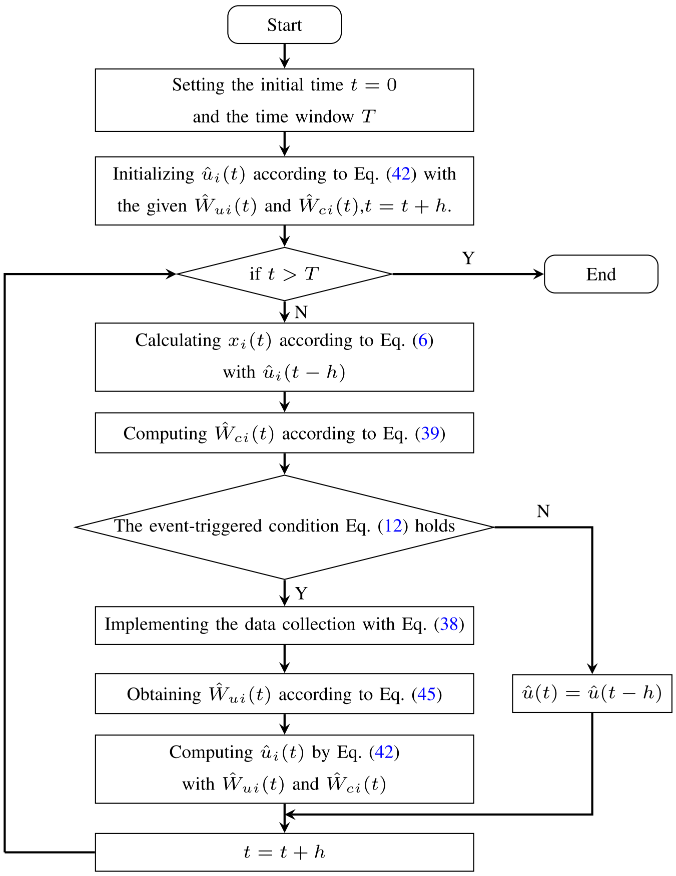

Fig. 1. The detailed workflow of RL-based algorithm.

Proof: Construct the following Lynapunov candidates for each nonlinear interconnected subsystem:

$$\begin{aligned} L_i &= L_{ia}(x_i) + L_{ib}(x_{i,k}) + L_{ic} + L_{iu} \\ &= V_i^*(x_i) + V_i^*(x_{i,k}) + \frac{1}{2}tr\{\tilde{W}_{ci}^T\tilde{W}_{ci}\} + \frac{1}{2}tr\{\tilde{W}_{ui}^T\tilde{W}_{ui}\}. \end{aligned} \tag{48}$$

As is mentioned above, S_i' includes two parts: continuous time systems and discrete time systems. Therefore, the analysis of stability is divided into two situations according to whether the event is triggered.

Situation I: when the event is not triggered, that is $t \in [t_k^i h, t_{k+1}^i h)$. Then it has:

$$\dot{L}_i = \dot{L}_{ia}(x_i) + \dot{L}_{ic}. \tag{49}$$

First of all, differentiating $L_i(x_i)$ in Eq. (49) with respect to time t based on the S_i': $\dot{x}_i(t) = f_i(x_i) + g_i(x_i)\hat{u}_i(x_{i,k})$. It can be obtained:

$$\dot{L}_{ia}(x_i) = (V_{x_i}^*)^T(f_i(x_i) + g_i(x_i)\hat{u}_i(x_{i,k})). \tag{50}$$

According to the Eqs. (22)-(23), it has:

$$\begin{aligned} \dot{L}_{ia}(x_i) \le\ & \alpha_i V_i^* - x_i^T Q_i x_i - \eta_i P_i^2(x_i) \\ & - 2\lambda_i(u_i^*(x_i))^T R_i \Gamma^{-1}(u_i^*(x_i)/\lambda_i) \\ & + (V_{x_i}^*)^T g_i(x_i)\hat{u}_i(x_{i,k}). \end{aligned} \tag{51}$$

Meanwhile, using the similar method [35], it can be obtained that:

$$\begin{aligned} \dot{L}_{ia}(x_i) \le\ & \alpha_i V_i^* - x_i^T Q_i x_i - \eta_i P_i^2(x_i) \\ & \times \int_{u_i^*(x_i)}^{\hat{u}_i(x_{i,k})} 2\lambda_i(\Gamma^{-1}(\upsilon/\lambda_i) + \mathcal{D}_i^*(x_i))^T R_i d\upsilon. \end{aligned} \tag{52}$$

Denoting $\upsilon = -\lambda_i\Gamma(\varpi)$, from Eq. (52), we can derive:

$$\begin{aligned} & \int_{u_i^*(x_i)}^{\hat{u}_i(x_{i,k})} 2\lambda_i(\Gamma^{-1}(\upsilon/\lambda_i) + \mathcal{D}_i^*(x_i))^T R_i d\upsilon \\ & \le \int_{\mathcal{D}_i^*(x_i)}^{\hat{\mathcal{D}}_i(x_{i,k})} 4\lambda_i^2(\upsilon - \mathcal{D}_i^*(x_{i,k}))^T R_i d\upsilon \\ & = 2\lambda_i^2||\mathcal{D}_i^*(x_i) - \hat{\mathcal{D}}_i(x_{i,k})||^2||R_i|| \\ & = 2\lambda_i^2||R_i||||\mathcal{D}_i^*(x_i) - \mathcal{D}_i^*(x_{i,k}) + \mathcal{D}_i^*(x_{i,k}) - \hat{\mathcal{D}}i(x_{i,k})||^2 \\ & \le 2\lambda_i^2||R_i||((1 + \frac{1}{\gamma_i})\mathcal{L}_i^2||z_{i,k}||^2 \\ & \quad + (1+\gamma_i)||\mathcal{D}_i^*(x_{i,k}) - \hat{\mathcal{D}}_i(x_{i,k})||^2), \end{aligned} \tag{53}$$

where $\gamma_i \in (0, 1)$.

Then it implies:

$$\begin{aligned} & ||\mathcal{D}_i^*(x_{i,k}) - \hat{\mathcal{D}}_i(x_{i,k})||^2 \\ & = ||\frac{1}{2\lambda_i}R_i^{-1}g_i^T(x_{i,k})\nabla\phi_{ci}^T(x_{i,k})\tilde{W}_{ci} \\ & \quad + \frac{1}{2\lambda_i}R_i^{-1}g_i^T(x_{i,k})\nabla\varepsilon_{ci}(x_{i,k})||^2 \\ & \le \frac{1+1/\gamma_i}{4\lambda_i^2}||R_i||^{-2}g_{im}^2\nabla\phi_{cim}^2||\tilde{W}_{ci}||^2 \\ & \quad + \frac{1+\gamma_i}{4\lambda_i^2}||R_i||^{-2}g_{im}^2\nabla\varepsilon_{cim}^2. \end{aligned} \tag{54}$$

Based on the Eqs. (51)-(54), $\dot{L}_{ia}(x_i)$ can be transformed into:

$$\begin{aligned} \dot{L}_{ia}(x_i) \le\ & \alpha_i V_i^* - x_i^T Q_i x_i - \eta_i P_i^2(x_i) \\ & + 2\lambda_i^2||R_i||(1 + \frac{1}{\gamma_i})\mathcal{L}_i^2||z_{i,k}||^2 \\ & + 2\lambda_i^2\frac{(1+\gamma_i)^2}{4\gamma_i\lambda_i^2}||R_i||^{-1}g_{im}^2\nabla\phi_{cim}^2||\tilde{W}_{ci}||^2 \\ & + 2\lambda_i^2\frac{(1+\gamma_i)^2}{4\lambda_i^2}||R_i||^{-1}g_{im}^2\nabla\varepsilon_{cim}^2. \end{aligned} \tag{55}$$

On the other hand, according to the Eqs. (39)-(40), $\dot{L}_{ic}$ can be rewritten as:

$$\begin{aligned} \dot{L}_{ic} =\ & -\iota_{ci}\tilde{W}_{ci}^T(\frac{\vartheta_i\vartheta_i^T}{(1+\vartheta_i^T\vartheta_i)^2} + \sum_{s=1}^{N_i(t)}\frac{\vartheta_{i,s}\vartheta_{i,s}^T e^{t_s^i h - t}}{(1+\vartheta_{i,s}^T\vartheta_{i,s})^2})\tilde{W}_{ci} \\ & + \frac{\iota_{ci}\tilde{W}_{ci}^T\vartheta_i}{(1+\vartheta_i^T\vartheta_i)^2}\varrho_{ci} + \sum_{s=1}^{N_i(t)}\frac{\iota_{ci}e^{t_s^i h - t}\tilde{W}_{ci}^T\vartheta_{i,s}}{(1+\vartheta_{i,s}^T\vartheta_{i,s})^2}\varrho_{ci,s}. \end{aligned} \tag{56}$$

Utilizing the Young's inequality, it yields that:

$$\dot{L}_{ic} \le -\frac{\iota_{ci}}{2}\tilde{W}_{ci}^T(\frac{\vartheta_i\vartheta_i^T}{(1+\vartheta_i^T\vartheta_i)^2} + \sum_{s=1}^{N_i(t)}\frac{\vartheta_{i,s}\vartheta_{i,s}^T e^{t_s^i h - t}}{(1+\vartheta_{i,s}^T\vartheta_{i,s})^2})\tilde{W}_{ci}$$

$$
+\frac{\iota_{ci}\varrho_{ci}^{T}\varrho_{ci}}{2(1+\vartheta_i^T\vartheta_i)^2}+\sum_{s=1}^{N_i(t)}\frac{\iota_{ci}\varrho_{ci,s}^{T}\varrho_{ci,s}e^{t_s^i h-t}}{2(1+\vartheta_{i,s}^T\vartheta_{i,s})^2}
\le -\frac{\iota_{ci}}{2}\lambda_{min}(\mathcal{H}_i)||\tilde{W}_{ci}||^2+\frac{\iota_{ci}(N_i(t)+1)}{2}\varrho_{im}^2, \quad (57)
$$

where $\mathcal{H}_i = \frac{\vartheta_i\vartheta_i^T}{(1+\vartheta_i^T\vartheta_i)^2}+\sum_{s=1}^{N_i(t)}\frac{\vartheta_{i,s}\vartheta_{i,s}^T e^{t_s^i h-t}}{(1+\vartheta_{i,s}^T\vartheta_{i,s})^2}$. Summarizing the analysis above and combining with the event-triggered mechanism (28), $\dot{L}_i$ can be transformed as:

$$
\begin{aligned}
\dot{L}_i &\le \alpha_i V_{im}^* + 2\lambda_i^2\frac{(1+\gamma_i)^2}{4\lambda_i^2}||R_i||^{-1}g_{im}^2\nabla\varepsilon_{cim}^2 \\
&\quad + \frac{\iota_{ci}(N_i(t)+1)}{2}\varrho_{im}^2 - (\lambda_{\min}(Q_i) - 2\lambda_i^2\mathcal{L}_i^2 \\
&\quad (1+1/\gamma_i)||R_i||\frac{\bar{\zeta}_i\lambda_{\max}(\Psi_i)}{\lambda_{\min}(\Psi_i)})||x_i||^2 - (\frac{\iota_{ci}}{2}\lambda_{min}(\mathcal{H}_i) \\
&\quad - 2\lambda_i^2\frac{(1+\gamma_i)^2}{4\gamma_i\lambda_i^2}||R_i||^{-1}g_{im}^2\nabla\phi_{cim}^2)||\tilde{W}_{ci}||^2 \\
&\le \tau_{i1} - \tau_{ix}||x_i||^2 - \tau_{iwc}||\tilde{W}_{ci}||^2, \quad (58)
\end{aligned}
$$

where

$$
\begin{cases}
\tau_{i1} = \alpha_i V_{im}^* + 2\lambda_i^2\frac{(1+\gamma_i)^2}{4\lambda_i^2}||R_i||^{-1}g_{im}^2\nabla\varepsilon_{cim}^2 \\
\quad +\frac{\iota_{ci}(N_i(t)+1)}{2}\varrho_{im}^2, \\
\tau_{ix} = \lambda_{\min}(Q_i) - 2\lambda_i^2\mathcal{L}_i^2(1+1/\gamma_i)||R_i||\frac{\bar{\zeta}_i\lambda_{\max}(\Psi_i)}{\lambda_{\min}(\Psi_i)}, \\
\tau_{iwc} = \frac{\iota_{ci}}{2}\lambda_{min}(\mathcal{H}_i) - 2\lambda_i^2\frac{(1+\gamma_i)^2}{4\gamma_i\lambda_i^2}||R_i||^{-1}g_{im}^2\nabla\phi_{cim}^2.
\end{cases} \quad (59)
$$

Then it implies that $\dot{L}_i < 0$ when $t \in [t_k^i h, t_{k+1}^i h)$, if condition (47) holds, $x_i \notin \mathcal{V}_{x_i}$ and $\tilde{W}_{ci} \notin \mathcal{V}_{\tilde{W}_{ci}}$, respectively, in which $\mathcal{V}_{x_i}$ and $\mathcal{V}_{\tilde{W}_{ci}}$ are defined as:

$$
\begin{cases}
\mathcal{V}_{x_i} = \left\{x_i : ||x_i|| \le \sqrt{\frac{\tau_{i1}}{\tau_{ix}}}\right\} \\
\mathcal{V}_{\tilde{W}_{ci}} = \left\{\tilde{W}_{ci} : ||\tilde{W}_{ci}|| \le \sqrt{\frac{\tau_{i1}}{\tau_{iwc}}}\right\}
\end{cases} \quad (60)
$$

According to the Definition 1, the UUB stability of ith subsystem x_i and weight vector estimation error of critic network $\tilde{W}_{ci}$ is guaranteed based on Lyapunov extension theorem. In addition, the ultimate bound of them are derived in Eq. (60).

Situation II: When the event is triggered, which means $t = t_{k+1}^i h$. The the difference of L_i with respect to t can be described as:

$$
\begin{aligned}
\Delta L_i &= V_i^*(x_{i,k}^+) - V_i^*(x_{i,k}) + V_i^*(x_{i,k+1}) - V_i^*(x_{i,k}) \\
&\quad + \frac{1}{2}tr\{\tilde{W}_{ci}^T(x_{i,k}^+)\tilde{W}_{ci}(x_{i,k}^+)\} - \frac{1}{2}tr\{\tilde{W}_{ci}^T(x_{i,k})\tilde{W}_{ci}(x_{i,k})\} \\
&\quad + \frac{1}{2}tr\{\tilde{W}_{ui}^T(x_{i,k}^+)\tilde{W}_{ui}(x_{i,k}^+)\} - \frac{1}{2}\{\tilde{W}_{ui}^T(x_{i,k})\tilde{W}_{ui}(x_{i,k})\},
\end{aligned} \quad (61)
$$

where $x_i^+(t) = \lim_{\kappa\to 0^+} x_i(t_k^i+\kappa)$.

Moreover, it implies that $\dot{L}_i < 0$ for $t \in [t_k^i h, t_{k+1}^i h)$ if $x_i \notin \mathcal{V}_{xi}$ and $\tilde{W}_{ci} \notin \mathcal{V}_{iwc}$ hold, which means that $L_{ia}(x_i) + L_{ic}$ is strictly monotonically decreasing over $t \in [t_k^i h, t_{k+1}^i h)$. That is to say:

$$
\begin{aligned}
L_{ia}(x_{i,k}) + L_{ic}(x_{i,k}) &> L_{ia}(x_i(t_k^i+\kappa)) + L_{ic}(t_k^i+\kappa), \\
&\qquad \kappa \in [0, (t_{k+1}^i - t_k^i)h).
\end{aligned} \quad (62)
$$

Let $\kappa \to 0^+$, then Eq. (62) can be transformed as:

$$
\begin{aligned}
L_{ia}(x_{i,k}) + L_{ic}(x_{i,k}) &> \lim_{\kappa\to 0^+} L_{ia}(x_i(t_k^i+\kappa)) + L_{ic}(t_k^i+\kappa) \\
&= L_{ia}(x_{i,k}^+) + L_{ic}(x_{i,k}^+).
\end{aligned} \quad (63)
$$

Furthermore, according to the UUB stability for x_i and $\tilde{W}_{ci}$ proved in **Situation I** when $t \in [t_k^i h, t_{k+1}^i h)$, it yields that

$$
V_i^*(x_{i,k+1}) \le V_i^*(x_{i,k}). \quad (64)
$$

Next, the term ΔL_{iu} can be dealt with:

$$
\begin{aligned}
\Delta L_{iu} &\le \frac{1}{2}tr\{\tilde{W}_{ui}^T(x_{i,k}^+)\tilde{W}_{ui}(x_{i,k}^+)\} \\
&\quad - \frac{1}{2}tr\{\tilde{W}_{ui}^T(x_{i,k})\tilde{W}_{ui}(x_{i,k})\} \\
&= -\iota_{ui}tr\{\tilde{W}_{ui}^T\phi_{u_i}(x_{i,k})\varepsilon_{ui}^T(x_{i,k}) \quad (65) \\
&\quad + \tilde{W}_{ui}^T\phi_{u_i}(x_{i,k})\phi_{u_i}^T(x_{i,k})\tilde{W}_{ui} \\
&\quad + \tilde{W}_{ui}^T\phi_{u_i}(x_{i,k})\lambda_i[\Gamma(\hat{\mathcal{D}}_i(x_{i,k})) - \Gamma(\mathcal{D}_i^*(x_{i,k}))]\} \\
&\quad + \frac{\iota_{ui}^2}{2}tr\{(\phi_{u_i}(x_{i,k}))^2[(\varepsilon_{ui}^T(x_{i,k}))^2 + (\phi_{u_i}^T(x_{i,k})\tilde{W}_{ui})^2 \\
&\quad + (\lambda_i[\Gamma(\hat{\mathcal{D}}_i(x_{i,k})) - \Gamma(\mathcal{D}_i^*(x_{i,k}))])^2 \\
&\quad + 2\phi_{u_i}^T(x_{i,k})\tilde{W}_{ui}\varepsilon_{ui}(x_{i,k}) \\
&\quad + 2\phi_{u_i}^T(x_{i,k})\tilde{W}_{ui}\lambda_i[\Gamma(\hat{\mathcal{D}}_i(x_{i,k})) - \Gamma(\mathcal{D}_i^*(x_{i,k}))] \\
&\quad + 2\varepsilon_{ui}^T(x_{i,k})\lambda_i[\Gamma(\hat{\mathcal{D}}_i(x_{i,k})) - \Gamma(\mathcal{D}_i^*(x_{i,k}))]]\}.
\end{aligned}
$$

Besides, considering the control constraint $|u_i| < \lambda_i$, then it has:

$$
tr\{\lambda_i|\Gamma(\hat{\mathcal{D}}_i(x_{i,k})) - \Gamma(\mathcal{D}_i^*(x_{i,k}))|\} < \Upsilon_{im}. \quad (66)
$$

Based on the property of function $tanh(\cdot)$, Young's inequality and Assumptions, then the Eq. (65) can be transformed into:

$$
\begin{aligned}
\Delta L_{iu} &\le \iota_{ui}(-\phi_{u_{im}}^2 + 1 + \frac{\iota_{ui}^2}{2}\phi_{uim}^4)||\tilde{W}_{ui}||^2 \\
&\quad + (\iota_{ui}^2\phi_{uim}^3\varepsilon_{uim} + \iota_{ui}^2\phi_{uim}^3\lambda_i\Upsilon_{im} \\
&\quad + \iota_{ui}^2\phi_{uim}^2\varepsilon_{uim}\lambda_i\Upsilon_{im})||\tilde{W}_{ui}|| \\
&\quad + (\frac{1}{2}\phi_{uim}^2\varepsilon_{uim}^2 + \frac{\iota_{ui}^2}{2}\phi_{uim}^2\varepsilon_{uim}^2 \\
&\quad + \frac{\iota_{ui}^2}{2}\lambda_i^2\Upsilon_{im}^2 + \frac{1}{2}\phi_{uim}^2\lambda_i^2\Upsilon_{im}^2) \\
&= -\rho_{i1}||\tilde{W}_{ui}||^2 + \rho_{i2}||\tilde{W}_{ui}|| + \rho_{i3},
\end{aligned} \quad (67)
$$

where

$$
\begin{cases}
\rho_{i1} = \phi_{u_{im}}^2 - 1 - \frac{\iota_{ui}^2}{2}\phi_{uim}^4 \\
\rho_{i2} = \iota_{ui}^2\phi_{uim}^3\varepsilon_{uim} + \iota_{ui}^2\phi_{uim}^3\lambda_i\Upsilon_{im} \\
\quad +\iota_{ui}^2\phi_{uim}^2\varepsilon_{uim}\lambda_i\Upsilon_{im} \\
\rho_{i3} = \frac{1}{2}\phi_{uim}^2\varepsilon_{uim}^2 + \frac{\iota_{ui}^2}{2}\phi_{uim}^2\varepsilon_{uim}^2 \\
\quad +\frac{\iota_{ui}^2}{2}\lambda_i^2\Upsilon_{im}^2 + \frac{1}{2}\phi_{uim}^2\lambda_i^2\Upsilon_{im}^2
\end{cases} \quad (68)
$$

Thus it can be concluded that $\Delta L_{iu} < 0$ when $\tilde{W}_{ui}$ lies in the set $\mathcal{V}_{\tilde{W}_{ui}}$:

$$\mathcal{V}_{\tilde{W}_{ui}} = \{\tilde{W}_{ui} : ||\tilde{W}_{ui}|| > \frac{\rho_{i2} + \sqrt{\rho_{i2}^2 + 4\rho_{i1}\rho_{i3}}}{2\rho_{i1}}\}. \quad (69)$$

According to the analysis above, the UUB stability of system states x_i, estimation error of W_{ci} and W_{ui} can be guaranteed based on the Lyapunov extension theorem. ■

IV. Simulation Result

In order to validate the effectiveness of the proposed RL-based optimal decentralized control policy, a simulation test for the NISs consisting two subsystems is conducted with the following parameters:

$$f_1(x_1) = \begin{bmatrix} x_{12} \\ -1.905sin(x_{11}) - 2.18sin(x_{12}) \end{bmatrix},$$
$$f_2(x_2) = \begin{bmatrix} 2.2x_{22} \\ -3.905sin(x_{21}) - 1.02sin(x_{22}) \end{bmatrix},$$
$$g_1(x_1) = \begin{bmatrix} -1.2 \\ 1.1 \end{bmatrix}, \quad g_2(x_2) = \begin{bmatrix} -1.8 \\ 1.2 \end{bmatrix},$$
$$h_1(x_1) = \begin{bmatrix} -0.5 \\ 0.4 \end{bmatrix}, \quad h_2(x_2) = \begin{bmatrix} 0.5 \\ -0.5 \end{bmatrix}.$$

with $x_1 = [x_{11}, x_{12}]^T$, $x_2 = [x_{21}, x_{22}]^T$ being the state vectors of the considered NISs.

Meanwhile, the relative parameters in value function are designed: $Q_1 = diag\{400, 600\}$, $Q_2 = diag\{500, 500\}$, $R_1 = 0.5$, $R_2 = 0.5$, $\eta_1 = 0.9$, $\eta_2 = 1.25$, $\alpha_1 = 0.36$, $\alpha_2 = 0.42$.

Besides, the upper constraints of control input is set as $\lambda_1 = 0.4$, $\lambda_2 = 0.4$, and the triggering matrix of AETM is designed as $\Psi_1 = diag\{0.4, 0.6\}$, $\Psi_2 = diag\{1.5, 0.6\}$. Moreover, the upper bound of adaptive parameter is given as $\underline{\zeta}_1 = 0.001$, $\underline{\zeta}_2 = 0.01$, $\bar{\zeta}_1 = 0.005$, $\bar{\zeta}_2 = 0.05$, $\varepsilon_1 = \varepsilon_2 = 0.02$.

Besides, the activation function in actor-critic network is designed as the following form:

$$\begin{cases} \phi_{ci} = \begin{bmatrix} x_{i1}^2 & x_{i1}x_{i2} & x_{i2}^2 \end{bmatrix}^T \\ \phi_{ui} = \begin{bmatrix} x_{i1}^3 & x_{i1}^2x_{i2} & x_{i1}x_{i2} & x_{i1}x_{i2}^2 & x_{i2}^3 \end{bmatrix}^T \end{cases} \quad i = 1, 2,$$

and the learning rate is given $\iota_{c1} = 0.01$, $\iota_{c2} = 0.03$, $\iota_{u1} = 0.02$, $\iota_{u2} = 0.02$.

In addition, the initial state of subsystems are given as $x_1 = \begin{bmatrix} 0.6 & -2.2 \end{bmatrix}^T$, $x_2 = \begin{bmatrix} -0.2 & 0.2 \end{bmatrix}^T$. Moreover, a small exploratory signal $o_i(t)$ is added to $\hat{u}_i$:

$$o_i(t) = 1000e^{-t}sin(t)^5cos(t) + sin(t)cos(t)^5, \quad i = 1, 2. \quad (70)$$

Then the simulation results are shown in Figs. 2-6.

As shown in the systems fluctuation in Fig. 2, it is easily seen that the RL-based optimal adaptive event-triggered decentralized control method can successfully stabilize NISs in the sense of UUB under control constraints. Moreover, both of actor and critic network vectors can converge to the certain value under proposed RL-based iterative algorithm, which are represented in Figs.4-5. Moreover, the actual control input value are depicted in Fig. 3 under control constraints. Besides, the event-triggered instants are shown in Fig. 6, which further validates the effectiveness of proposed optimal decentralized control strategies.

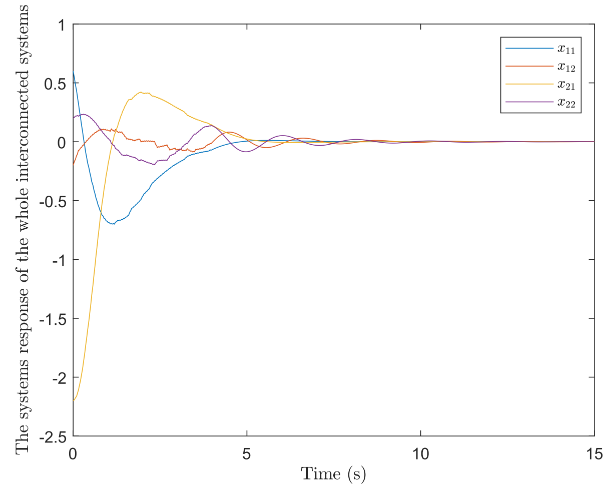

Fig. 2. The states response of whole NISs.

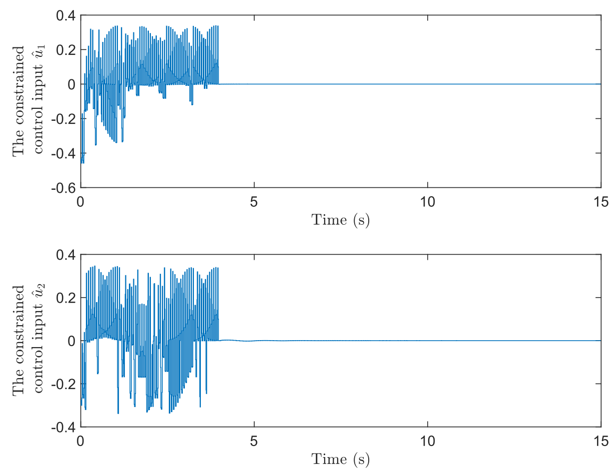

Fig. 3. The values of constrained control input $\hat{u}_1$ and $\hat{u}_2$.

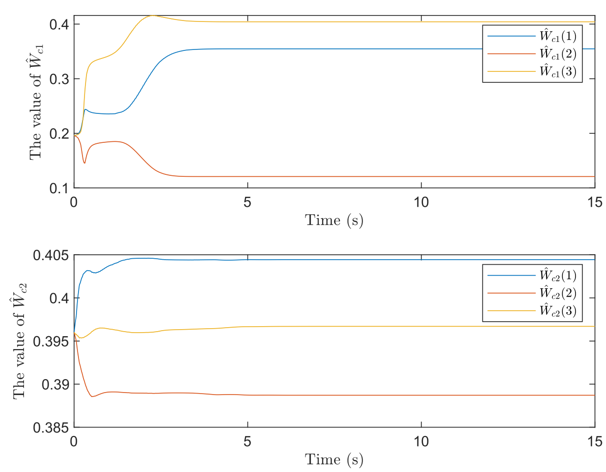

Fig. 4. The values of $\hat{W}_{c1}$ and $\hat{W}_{c2}$.

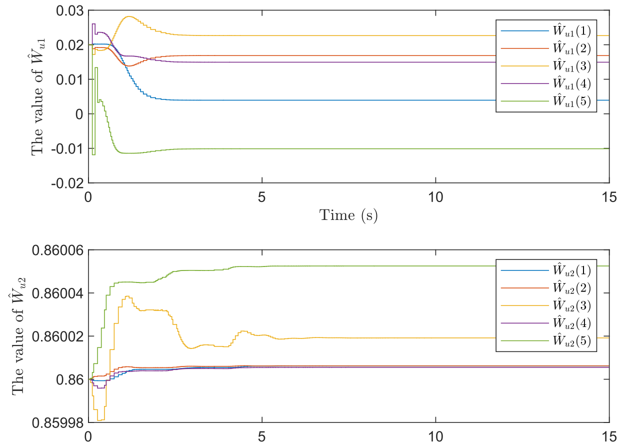

Fig. 5. The values of $\hat{W}_{u1}$ and $\hat{W}_{u2}$.

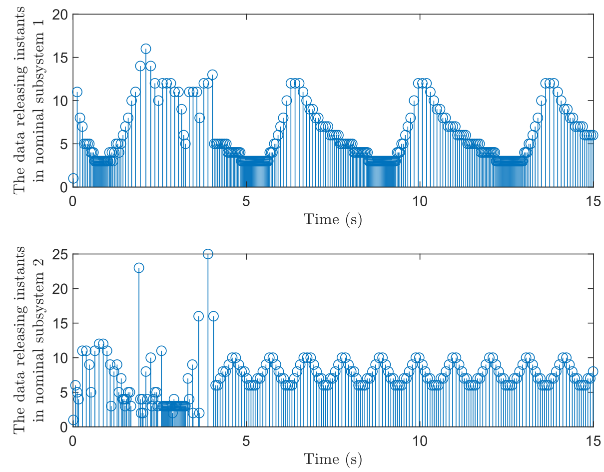

Fig. 6. The data releasing instants in nominal subsystem 1 and subsystem 2.

To validate the advantages of the AETM-based RL algorithm proposed in this paper, we then compare some results obtained under the designed algorithm and relative algorithm proposed in the reference [35]. The event-triggered mechanism in [35] is derived during the operation of proving the control performance of NISs, which is designed to assist the stabilization of the considered NISs. Whereas the event-triggered mechanism in this paper is pre-designed to save the network bandwidth. Specifically, for the optimal decentralized control problem of nonlinear NISs with control and communication constraints in this paper, the event-triggering condition designed according to [35] is:

$$||z_{i,k}||^2 \leq \frac{(1-q_i^2)\lambda_{min}(Q_i)||x_i||^2}{2\lambda_i^2\mathcal{L}_i^2||R_i||(1+1/\gamma_i)}, \tag{71}$$

where $q_i \in (0,1)$ is a given constant and $\mathcal{L}_i > 0$ is the pre-given parameter which are then set to be $q_i = 0.34$ and $\mathcal{L}_i = 21$ in the simulation. With the triggering condition, the ultimate bounds of the system state x_i and approximate weight error $\tilde{W}_{ci}$ are:

$$\begin{cases} \mathcal{V}_{x_i} = \left\{x_i : ||x_i|| \leq \sqrt{\dfrac{\tau_{i1a}}{p_i\lambda_{\min}(Q_i)}}\right\}, \\ \mathcal{V}_{\tilde{W}_{ci}} = \left\{\tilde{W}_{ci} : ||\tilde{W}_{ci}|| \leq \sqrt{\dfrac{\tau_{i1a}}{\tau_{iwca}}}\right\}, \end{cases} \tag{72}$$

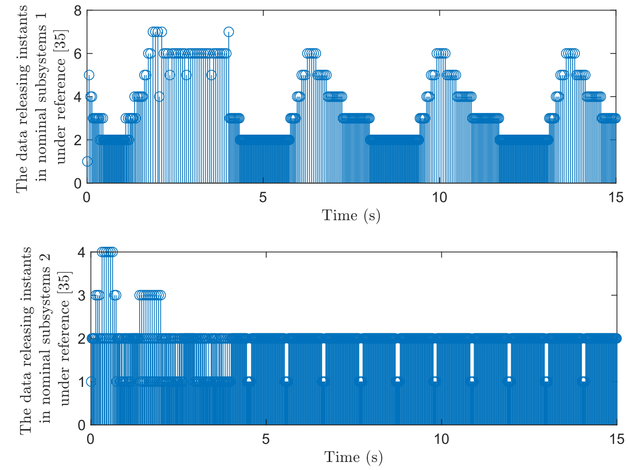

Fig. 7. The event-triggered instants of nominal subsystem 1 and subsystem 2 under the algorithm proposed in [35].

where

$$\begin{cases} \tau_{i1a} = \alpha_i V_{im}^* + 2\lambda_i^2\dfrac{(1+\gamma_i)^2}{4\lambda_i^2}||R_i||^{-1}g_{im}^2\nabla\varepsilon_{cim}^2 \\ \quad + \dfrac{\iota_{ci}(N_i(t)+1)}{2}\varrho_{im}^2, \\ \tau_{iwca} = \dfrac{\iota_{ci}}{2}\lambda_{min}(\mathcal{H}_i) - 2\lambda_i^2\dfrac{(1+\gamma_i)^2}{4\gamma_i\lambda_i^2}||R_i||^{-1}g_{im}^2\nabla\phi_{cim}^2. \end{cases} \tag{73}$$

In Fig. 7, the triggering instants for the subsystems are presented and the system stability can also be ensured (owing to page limitation, the system states operations are omitted). From Fig. 6 and Fig. 7, it can be found that less triggering instants are generated with the algorithm proposed in this paper (actually, the numbers of the triggering instants in this paper are 277 and 214, with the triggering instants are 492 and 834 under the referred algorithm [35] within 15s, respectively.). Obviously, under the AETM proposed in this paper, the bandwidth pressure can be effectively alleviated.

V. Conclusion

This paper has addressed the AETM-based decentralized optimal control problem for a class of NISs with control constraints. Firstly, the decentralized optimal problem was transformed into the optimal control for corresponding relative nominal subsystems and the HJI equations were established via the optimal control theory. Meanwhile, the UUB stability of NISs was proved under optimal control strategies. In order to solve the HJI equations, an actor-critic network structure was utilized to obtain the numerical solutions of the HJI equations and the tuning law of weight vectors were designed with the collocation of historical system sates. By means of Lyapunov stability theorem and optimal control theory, the estimation error of weight vectors in actor-critic network and system states were guaranteed to be stable in the sense of UUB under proposed RL-based control strategies. Eventually, a simulation test was conducted to further validate the effectiveness of control methods.

REFERENCES

[1] N. Zhao, P. Shi, W. Xing, and C. P. Lim, "Event-triggered control for networked systems under denial of service attacks and applications," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 69, no. 2, pp. 811–820, Feb. 2022.

[2] X. Jin, S. Lü, and J. Yu, "Adaptive NN-based consensus for a class of nonlinear multiagent systems with actuator faults and faulty networks," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 33, no. 8, pp. 3474–3486, Aug. 2022.

[3] M. Bahreini, J. Zarei, R. Razavi-Far, and M. Saif, "Robust and reliable output feedback control for uncertain networked control systems against actuator faults," *IEEE Trans. Syst. Man, Cybern. Syst.*, vol. 52, no. 4, pp. 2555–2564, Apr. 2022.

[4] B. Yan, P. Shi, and C.-C. Lim, "Robust formation control for nonlinear heterogeneous multiagent systems based on adaptive event-triggered strategy," *IEEE Trans. Autom. Sci. Eng.*, vol. 19, no. 4, pp. 2788–2800, Oct. 2022.

[5] N. Rong and Z. Wang, "Event-based fixed-time control for interconnected systems with discontinuous interactions," *IEEE Trans. Syst. Man, Cybern. Syst.*, vol. 52, no. 8, pp. 4925–4936, Aug. 2022.

[6] H. Zhang, C. Liu, H. Su, and K. Zhang, "Echo state network-based decentralized control of continuous-time nonlinear large-scale interconnected systems," *IEEE Trans. Syst. Man, Cybern. Syst.*, vol. 51, no. 10, pp. 6293–6303, Oct. 2021.

[7] H. Sun, L. Hou, G. Zong, and X. Yu, "Adaptive decentralized neural network tracking control for uncertain interconnected nonlinear systems with input quantization and time delay," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 31, no. 4, pp. 1401–1409, Apr. 2020.

[8] Y. Xu, H. Yang, B. Jiang, and M. M. Polycarpou, "Distributed optimal fault estimation and fault-tolerant control for interconnected systems: A Stackelberg differential graphical game approach," *IEEE Trans. Autom. Control*, vol. 67, no. 2, pp. 926–933, Feb. 2022.

[9] Y. Xu, H. Yang, and B. Jiang, "Fault-tolerant control of multilayer interconnected nonlinear systems: An inclusion principle approach," *IEEE Trans. Syst. Man, Cybern. Syst.*, vol. 51, no. 4, pp. 2403–2414, Apr. 2021.

[10] D. Yang, T. Li, X. Xie, and H. Zhang, "Event-triggered integral sliding-mode control for nonlinear constrained-input systems with disturbances via adaptive dynamic programming," *IEEE Trans. Syst. Man, Cybern. Syst.*, vol. 50, no. 11, pp. 4086–4096, Nov. 2020.

[11] L. N. Tan, "Event-triggered distributed H_∞ constrained control of physically interconnected large-scale partially unknown strict-feedback systems," *IEEE Trans. Syst. Man, Cybern. Syst.*, vol. 51, no. 4, pp. 2444–2456, Apr. 2021.

[12] X. Yang, Y. Zhou, N. Dong, and Q. Wei, "Adaptive critics for decentralized stabilization of constrained-input nonlinear interconnected systems," *IEEE Trans. Syst. Man, Cybern. Syst.*, vol. 52, no. 7, pp. 4187–4199, Jul. 2022.

[13] Y. Chen, L. Pan, and Z. Li, "Design of optimal supervisors for the enforcement of nonlinear constraints on Petri nets," *IEEE Trans. Autom. Sci. Eng.*, vol. 20, no. 1, pp. 611–623, Jan. 2023.

[14] M. Wang, L. Huang, and C. Yang, "NN-based adaptive tracking control of discrete-time nonlinear systems with actuator saturation and event-triggering protocol," *IEEE Trans. Syst. Man, Cybern. Syst.*, vol. 51, no. 12, pp. 7613–7621, Dec. 2021.

[15] Y. Dong and Z. Lin, "An event-triggered observer and its applications in cooperative control of multiagent systems," *IEEE Trans. Autom. Control*, vol. 67, no. 7, pp. 3647–3654, Jul. 2022.

[16] Y. Yang et al., "An event-triggered hybrid system model for cascading failure in power grid," *IEEE Trans. Autom. Sci. Eng.*, vol. 19, no. 3, pp. 1312–1325, Jul. 2022.

[17] X.-G. Guo, D.-Y. Zhang, J.-L. Wang, J. H. Park, and L. Guo, "Observer-based event-triggered composite anti-disturbance control for multi-agent systems under multiple disturbances and stochastic FDIAs," *IEEE Trans. Autom. Sci. Eng.*, vol. 20, no. 1, pp. 528–540, Jan. 2023.

[18] 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.

[19] S. Zhu, E. Tian, D. Xu, and J. Liu, "An adaptive torus-event-based controller design for networked T-S fuzzy systems under deception attacks," *Int. J. Robust Nonlinear Control*, vol. 32, no. 6, pp. 3425–3441, Apr. 2022.

[20] X. Li, Z. Sun, Y. Tang, and H. R. Karimi, "Adaptive event-triggered consensus of multiagent systems on directed graphs," *IEEE Trans. Autom. Control*, vol. 66, no. 4, pp. 1670–1685, Apr. 2021.

[21] J. Liu, Y. Wang, J. Cao, D. Yue, and X. Xie, "Secure adaptive-event-triggered filter design with input constraint and hybrid cyber attack," *IEEE Trans. Cybern.*, vol. 51, no. 8, pp. 4000–4010, Aug. 2021.

[22] Y. Xu and Z.-G. Wu, "Distributed adaptive event-triggered fault-tolerant synchronization for multiagent systems," *IEEE Trans. Ind. Electron.*, vol. 68, no. 2, pp. 1537–1547, Feb. 2021.

[23] Z. Li, J. Yan, W. Yu, and J. Qiu, "Adaptive event-triggered control for unknown second-order nonlinear multiagent systems," *IEEE Trans. Cybern.*, vol. 51, no. 12, pp. 6131–6140, Dec. 2021.

[24] H. Li, Y. Liu, and F. Li, "Adaptive event-triggered output feedback for nonlinear systems with unknown polynomial-of-output growth rate," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 69, no. 5, pp. 2179–2192, May 2022.

[25] C. Sun, X. Li, and Y. Sun, "A parallel framework of adaptive dynamic programming algorithm with off-policy learning," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 8, pp. 3578–3587, Aug. 2021.

[26] F. Hagebring, A. Farooqui, M. Fabian, and B. Lennartson, "On optimization of automation systems: Integrating modular learning and optimization," *IEEE Trans. Autom. Sci. Eng.*, vol. 19, no. 3, pp. 1662–1674, Jul. 2022.

[27] Y. Qian, D. Hu, Y. Chen, and Y. Fang, "Programming-based optimal learning sliding mode control for cooperative dual ship-mounted cranes against unmatched external disturbances," *IEEE Trans. Autom. Sci. Eng.*, vol. 20, no. 2, pp. 969–980, Apr. 2023.

[28] Z. Ji, C. Chen, J. He, S. Zhu, and X. Guan, "Learning-based edge sensing and control co-design for industrial cyber–physical system," *IEEE Trans. Autom. Sci. Eng.*, vol. 20, no. 1, pp. 59–73, Jan. 2023.

[29] Q. Wei, R. Song, Z. Liao, B. Li, and F. L. Lewis, "Discrete-time impulsive adaptive dynamic programming," *IEEE Trans. Cybern.*, vol. 50, no. 10, pp. 4293–4306, Oct. 2020.

[30] 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.

[31] 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.

[32] V. Narayanan, H. Modares, S. Jagannathan, and F. L. Lewis, "Event-driven off-policy reinforcement learning for control of interconnected systems," *IEEE Trans. Cybern.*, vol. 52, no. 3, pp. 1936–1946, Mar. 2022.

[33] R. Song, F. L. Lewis, Q. Wei, and H. Zhang, "Off-policy actor-critic structure for optimal control of unknown systems with disturbances," *IEEE Trans. Cybern.*, vol. 46, no. 5, pp. 1041–1050, May 2016.

[34] X. Wang, Q. Wang, and C. Sun, "Prescribed performance fault-tolerant control for uncertain nonlinear MIMO system using Actor–Critic learning structure," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 33, no. 9, pp. 4479–4490, Sep. 2022.

[35] X. Yang and H. He, "Decentralized event-triggered control for a class of nonlinear-interconnected systems using reinforcement learning," *IEEE Trans. Cybern.*, vol. 51, no. 2, pp. 635–648, Feb. 2021.

[36] 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.

[37] K. G. Vamvoudakis, "Event-triggered optimal adaptive control algorithm for continuous-time nonlinear systems," *IEEE/CAA J. Autom. Sinica*, vol. 1, no. 3, pp. 282–293, Jul. 2014.

[38] V. Narayanan, A. Sahoo, S. Jagannathan, and K. George, "Approximate optimal distributed control of nonlinear interconnected systems using event-triggered nonzero-sum games," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 30, no. 5, pp. 1512–1522, May 2019.

[39] F. Lin and R. D. Brandt, "An optimal control approach to robust control of robot manipulators," *IEEE Trans. Robot. Autom.*, vol. 14, no. 1, pp. 69–77, Feb. 1998.

[40] H. Modares and F. L. Lewis, "Optimal tracking control of nonlinear partially-unknown constrained-input systems using integral reinforcement learning," *Automatica*, vol. 50, no. 7, pp. 1780–1792, Jul. 2014.

[41] H. Modares, F. L. Lewis, and M. N. Sistani, "Online solution of nonquadratic two-player zero-sum games arising in the H_∞ control of constrained input systems," *Int. J. Adapt. Control Signal Process.*, vol. 28, nos. 3–5, pp. 232–254, Mar. 2014.

[42] X. Yang and H.-B. He, "Event-driven H_∞ constrained control using adaptive critic learning," *IEEE Trans. Cybern.*, vol. 51, no. 10, pp. 4860–4872, Oct. 2021.

[43] A. Saberi, "On optimality of decentralized control for a class of nonlinear interconnected systems," *Automatica*, vol. 24, no. 1, pp. 101–104, Jan. 1988.

[44] F. Lin, *Robust Control Design: An Optimal Control Approach*. Hoboken, NJ, USA: Wiley, 2007, pp. 101–131.

[45] X. Xie, S. Li, and B. Xu, "Adaptive event-triggered H_∞ fuzzy filtering for interval type-2 T–S fuzzy-model-based networked control systems with asynchronously and imperfectly matched membership functions," *J. Franklin Inst.*, vol. 356, no. 18, pp. 11760–11791, Dec. 2019.

[46] A.-K. Murad, H. Jie, and L.-L. Frank, *Nonlinear H_2/H_∞ Constrained FeedBack Control: A Practical Design Approach Using Neural Networks*. London, U.K.: Springer, 2006.

[47] 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, May 2019.

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

He was a Post-Doctoral Research Associate with the School of Automation, Southeast University, Nanjing, China, from December 2013 to June 2016. He was a Visiting Researcher/Scholar with the Department of Mechanical Engineering, The University of Hong Kong, Hong Kong, from October 2016 to October 2017. He was a Visiting Scholar with the Department of Electrical Engineering, Yeungnam University, Gyeongsan, South Korea, from November 2017 to January 2018. From June 2011 to May 2023, he was an Associate Professor and then a Professor with the Nanjing University of Finance and Economics, Nanjing. In June 2023, he joined the Nanjing University of Information Science and Technology, Nanjing, where he is currently a Professor with the School of Computer Science. His research interests include networked control systems, complex dynamical networks, and time delay systems.

Nan Zhang received the B.S. degree in computer science from the Nanjing University of Finance and Economics, Nanjing, China, in 2020, where he is currently pursuing the M.Sc. degree with the College of Information Engineering. His research interests include networked control systems and learning-based optimal control methods.

Lijuan Zha received the Ph.D. degree in control science and engineering from Donghua University, Shanghai, China, in 2018.

She was a Post-Doctoral Research Associate with the School of Mathematics, Southeast University, Nanjing, China. She is currently an Associate Professor with the Nanjing University of Finance and Economics, Nanjing. Her current research interests include networked control systems, neural networks, and complex dynamical 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.

From 2010 to 2014, he was a Senior Engineer with Metallurgical Corporation of China Ltd., Beijing, China. He is currently a Professor with the Institute of Advanced Technology, Nanjing University of Posts and Telecommunications, Nanjing, China. His research interests include fuzzy modeling and control synthesis, state estimations, optimization in process industries, and intelligent optimization algorithms. He serves as an Associate Editor for the *International Journal of Fuzzy Systems* and *International Journal of Control, Automation, and Systems*.

Engang Tian 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 Post-Doctoral 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, where he is currently a Professor with the School of Optical-Electrical and Computer Engineering. He has published more than 100 articles in refereed international journals. His research interests include networked control systems, cyber attack, and nonlinear stochastic control and filtering.