Intelligent Control in Asymmetric Decision-Making: An Event-Triggered RL Approach for Mismatched Uncertainties

Xiangnan Zhong, *Member, IEEE*, and Zhen Ni, *Senior Member, IEEE*

***Abstract*—Artificial intelligence (AI)-based multiplayer systems have attracted increasing attention across diverse fields. While most research focuses on simultaneous-move multiplayer games to achieve Nash equilibrium, there are complex applications that involve hierarchical decision-making, where certain players act before others. This power asymmetry increases the complexity of strategic interactions, especially in the presence of mismatched uncertainties that can compromise data reliability and decision-making. To this end, this article develops a novel event-triggered reinforcement learning (RL) approach for hierarchical multiplayer systems with mismatched uncertainties. Specifically, by establishing an auxiliary augment system and designing appropriate cost functions for the high-level leader and low-level followers, we reformulate the hierarchical robust control problem as an optimization task within the Stackelberg–Nash game framework. Furthermore, an event-triggered scheme is designed to reduce the computational overhead and a neural-RL-based method is developed to automatically learn the event-triggered control policies for hierarchical players. Theoretical analyses are conducted to 1) demonstrate the stability preservation of the designed robust-optimal transformation; 2) verify the achievement of Stackelberg–Nash equilibrium under the developed event-triggered policies; and 3) guarantee the boundedness of the impulsive closed-loop system. Finally, the simulation studies validate the effectiveness of the developed method.**

***Index Terms*—Asymmetric decision-making, event-triggered control, hierarchical optimization, mismatched uncertainties, neural networks, reinforcement learning (RL).**

I. Introduction

ARTIFICIAL intelligence (AI) has achieved growing success in networked systems where multiple intelligent agents collaborate or compete to achieve complex objectives [1], [2], [3], [4]. The applications span various domains, such as networked microgrid energy management [5] and connected intelligent transportation [6]. Among the current AI techniques, reinforcement learning (RL) has emerged as one of the most powerful tools to autonomously learn decision-making policies through trial and error [7], [8], [9], [10]. By continuously improving strategies based on interactions with the environment and other players, RL has revolutionized the way how intelligent agents navigate complex decision-making and multiplayer interactions [11], [12], [13], [14], [15]. These advancements enable agents to dynamically adapt their strategies based on real-time feedback, leading to enhanced performance in complex environments [16], [17], [18], [19], [20].

Despite the success, most of the existing studies focus on synchronized decision-making with Nash equilibrium, where all participants typically share an equivalent status and operate on the same level. As a result, no agent can unilaterally gain an advantage by deviating from the equilibrium strategy. However, these studies overlook the sequential order of policy execution, where one controller may have the authority to act with priority, while the others attempt to respond with appropriate counterstrategies. There are many practical applications in the real-world. For instance, in energy management and demand reduction applications [21], utility providers often enforce temporary reductions in electricity usage during peak hours to prevent excessive expenditures on high-priced power. In response, local businesses and households modify their energy consumption patterns to meet their own needs while adhering to utility guidelines. Such hierarchical optimization is also commonly seen in smart grid [22], traffic networks [23], [24], and autonomous driving [25], where a dominant player holds an advantage to determine the strategies first, while other subordinate players respond based on the dominant's decisions. These hierarchical systems introduce power asymmetries and sequential decision-making, which present challenges go beyond the standard Nash equilibrium framework. Effectively addressing these complexities requires tailored methods.

Fortunately, the Stackelberg game provides valuable insights into asymmetric decision-making. In the Stackelberg game, two players with distinct hierarchical roles are involved: 1) the high-level leader and 2) the low-level follower [26], [27], [28], [29]. The leader, serving as the dominant player, optimizes its performance and announces a policy first, while the follower responds optimally. Developing learning-based control algorithms for such system is more challenging than the simultaneous-move multiplayer systems due to the information asymmetry and player interdependencies. In [27],

Received 6 April 2025; accepted 21 June 2025. Date of publication 10 July 2025; date of current version 18 September 2025. This work was supported in part by the National Science Foundation under Grant 2047010, Grant 2047064, Grant 1947419, Grant 2117822, and Grant 2205205. The work of Zhen Ni was supported in part by the Department of Transportation under Grant 69A3552348304. This article was recommended by Associate Editor D. Zhao. *(Corresponding author: Xiangnan Zhong.)*

The authors are with the Department of Electrical Engineering and Computer Science, Florida Atlantic University, Boca Raton, FL 33431 USA (e-mail: xzhong@fau.edu; zhenni@fau.edu).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TSMC.2025.3583066.

Digital Object Identifier 10.1109/TSMC.2025.3583066

2168-2216 © 2025 IEEE. All rights reserved, including rights for text and data mining, and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission.
See https://www.ieee.org/publications/rights/index.html for more information.
Authorized licensed use limited to: Florida Atlantic University. Downloaded on February 18,2026 at 16:15:36 UTC from IEEE Xplore. Restrictions apply.

an adaptive learning framework was developed based on RL techniques for a two-player linear Stackelberg differential game. The authors ensured that the resulting policies of the leader and the follower constituted a Stackelberg equilibrium. The nonlinear hierarchical control problem for two-player input-affine systems was investigated in [30]. An adaptive critic algorithm was established to derive the Stackelberg equilibrium policy in such a problem. This mechanism has also been extended to scenarios involving one leader and multiple followers, where the decision-making scheme incorporates both hierarchy and simultaneity [31], [32]. Depending on the different assumptions of multiple followers, Mukaidani and Xu [33] designed the Pareto-based and Nash-based Stackelberg strategies respectively. In [34], the RL algorithm, incorporating a new form of value function, was designed from a Stackelberg–Nash perspective for linear hierarchical system with one leader and multiple followers. This structure was extended to nonlinear hierarchical control problem in [35]. A robust approach was developed in [36] by formulating the disturbed hierarchical system as a zero-sum game. Then, RL techniques were applied to achieve Stackelberg–Nash–Saddle equilibrium.

In the above research studies, the intelligent control approaches were designed based on the periodic learning scheme with time-triggered mechanism. However, due to the coupled input interactions and asymmetric decision-making dynamics, the hierarchical multiplayer systems usually exhibit extensive computational burden. This leads to a notably high computational load during the learning process, and may impact the learning performance in resource-constrained environments, such as limited computation bandwidth. To solve this problem, the event-triggered mechanism with intermittent control feedback was designed. By updating the control policy only when some specific conditions are violated, the event-triggered mechanism can effectively reduce the computation and communication burden [37], [38], [39], [40], [41], [42], [43]. In [44], an event-triggered optimal regulation was developed for the multiagent synchronization problem to reduce the controller updates and also maintain stability and optimality. Recently, this mechanism was introduced in the multiplayer Stackelberg–Nash games with matched uncertainty in [45] and has demonstrated its effectiveness. However, few studies address mismatched uncertainties, which are more prevalent and often span a broader range, for hierarchical multiplayer systems. This is critical in complex, unpredictable, and resource-constrained environment.

Motivated by the above observations, this article develops an event-triggered RL approach for mismatched uncertain hierarchical multiplayer systems. The major contributions can be summarized as follows.

1) In this article, we design a novel framework to reduce the computational complexity and handle the mismatched uncertainties for a class of systems characterized by asymmetric decision-making. Different from the conventional multiplayer systems [12], [13], [17], [19], where all players act simultaneously, this type of system involves a distinct power hierarchy that shapes the strategic interactions within the system. This hierarchical structure creates interdependence and couplings in the overall dynamics, increasing the complexity of control design as it necessitates solving the coupled Hamilton–Jacobi (HJ) equations.
2) By developing an auxiliary augment system with appropriate cost functions, we transform this hierarchical robust control problem into an optimization task under the framework of Stackelberg–Nash game to facilitate the learning-based control process. The theoretical analysis is provided to demonstrate that the solution obtained from the transformed system ensures the stability of original robust problem. Comparing with the related work [34], [35], [36], [42], [45], this article addresses mismatched uncertainties in a hierarchical decision-making environment, which is more demanding due to the greater difficulties in predicting system dynamics and enhanced interdependence.
3) We design an event-triggered control scheme tailored specifically for hierarchical leader-follower interactions. We demonstrate that the resulting event-triggered control policies can achieve Stackelberg–Nash equilibrium. Different from the existing event-triggered approaches for multiagent systems [38], [43], [44] that focus on simultaneous decision-making, this design ensures that the leader's decisions optimally influence the followers while reducing unnecessary control updates. The convergence analysis becomes more challenging due to the existence of couplings among players.
4) A neural-RL-based method is developed to automatically learn the event-triggered control policies for the high-level leader and low-level followers. Detailed theoretical studies are conducted to guarantee the boundedness of the impulsive closed-loop system in both the continuous and jump dynamic phases. These analyses provide rigorous guarantees on system stability and performance, validating that the proposed control policies effectively manage the complexities introduced by the hierarchical and event-triggered nature of the problem.

The reminder of this article is organized as follows. Section II formulates the robust hierarchical control problem with asymmetric decision-making into the framework of Stackelberg–Nash game. The event-triggered RL control design and neural network implementation are provided in Section III with stability guarantees. In Section IV, the simulation studies are shown to demonstrate the effectiveness of the proposed approach. Finally, Section V concludes this work.

II. Formulation of Mismatched Uncertainties in Hierarchical Multiplayer Systems

Consider the continuous-time hierarchical system with $N+1$ players $\mathcal{P} = \{0, 1, 2, \ldots, N\}$, where player 0 is the leader and other players $\mathcal{F} = \{1, 2, \ldots, N\}$ are the followers. This system involves an asymmetric decision-making process where the leader holds a dominant position and establishes the policy first, while the followers share equal status and respond to the

Authorized licensed use limited to: Florida Atlantic University. Downloaded on February 18,2026 at 16:15:36 UTC from IEEE Xplore. Restrictions apply.

leader's decisions. The system function is given as

$$\dot{x} = f(x) + g_0(x)u_0 + k_0(x)d_0(x) + \sum_{i=1}^{N} g_i(x)u_i + \sum_{i=1}^{N} k_i(x)d_i(x) \tag{1}$$

where $x \in \mathbb{R}^n$ is the state vector, $u_0 \in \mathbb{R}^{m_0}$ and $u_i \in \mathbb{R}^{m_i}$, $i \in \mathcal{F}$ are the policies controlled by the leader and ith follower, respectively, $d_0(x) \in \mathbb{R}^{l_0}$ and $d_i(x) \in \mathbb{R}^{l_i}$, $i \in \mathcal{F}$ are the unknown uncertainties applied on leader and ith follower, respectively. $f(x) \in \mathbb{R}^n$ is the system drift dynamics, $g_0(x) \in \mathbb{R}^{n\times m_0}$, $g_i(x) \in \mathbb{R}^{n\times m_i}$, $k_0(x) \in \mathbb{R}^{n\times l_0}$, and $k_i(x) \in \mathbb{R}^{n*l_i}$ are local input and disturbance dynamics for leader and ith follower, respectively. In this article, we consider the mismatched uncertainties which follows $g_0(x) \neq k_0(x)$ if $m_0 = l_0$, and $g_i(x) \neq k_i(x)$ if $m_i = l_i$. Assume that all the dynamic functions are Lipschitz continuous with $f(0) = 0$.

It is worth noting that the hierarchical multiplayer system (1) considered in this article is substantially different from the conventional simultaneous-move multiagent systems or zero-sum games in [17], [19]. Specifically, in this article, we focus on the asymmetric decision-making dynamics within a hierarchy, where the leader holds a distinct advantage in determining its policy before the followers responses. In contrast, conventional multiagent systems operate on a flat structure where decisions are made simultaneously for all players. Furthermore, the state x in system (1) is influenced by the policies of all the players, which leads to interdependence and couplings between their decisions. Conversely, each player in the conventional multiagent systems operates with its own state, and each agent's state evolves independently based on its policy. Due to these properties, the control design for hierarchical multiplayer system (1) is more complex and challenging.

Assumption 1: The unknown uncertainty $d_j(x)$, $j \in \mathcal{P}$, is bounded as $||d_j(x)|| \leq c_j\sigma_j(||x||)$, where $c_j \geq 0$ is the constant and $\sigma_j(\cdot)$ is the class $\mathcal{K}$ function. Besides, $d_j(0) = 0$ and $\sigma_j(0) = 0$.

Note that physical constraints exist in systems, limiting the magnitude and extent of uncertainties. Therefore, this assumption is reasonable and aligns with the inherent physical limitations of the systems. Define $\mathscr{D}_d(x) = ||\sum_{j=0}^{N} k_j(x)c_j\sigma_j(||x||)||$ as the upper bound for the overall uncertainties.

Assumption 2: The system (1) is controllable with $f(0) = 0$. Moreover, $\text{rank}(g_j(x)) = m_j (m_j < n)$ and $g_j^T(x)k_j(x) = 0$ for the jth player.

The assumption $g_j^T(x)k_j(x) = 0$ is a strict constraint that may exclude certain nonlinear systems with uncertainties. However, this condition helps avoid more restrictive assumptions, such as the boundedness of the Moore–Penrose pseudoinverse of $g_j(x)$, which can be computationally expensive, particularly in high-dimensional systems [39]. Therefore, to facilitate the discussion, this article considers $g_j^T(x)k_j(x) = 0$ for each player j, $j \in \mathcal{P}$.

The goal of this article is to develop a set of intermittent control policies based on the event-triggered RL scheme for hierarchical multiplayer system (1) to achieve asymptotic stability with reduced communication burden. However, due to the existence of unknown uncertainties, the communication data we rely on for learning cannot be trusted. Therefore, we transform this robust control problem into an optimal stabilization design with the following auxiliary nominal plant

$$\dot{x} = f(x) + g_0(x)u_0 + \big(I_n - g_0(x)g_0^+(x)\big)k_0(x)v_0 + \sum_{i=1}^{N} g_i(x)u_i + \sum_{i=1}^{N}\big(I_n - g_i(x)g_i^+(x)\big)k_i(x)v_i \tag{2}$$

where $g_0^+(x)$ and $g_i^+(x)$ are the Moore–Penrose pseudoinverse matrices of $g_0(x)$ and $g_i(x)$, respectively, $v_0 \in \mathbb{R}^{l_0}$ and $v_i \in \mathbb{R}^{l_i}$ are the auxiliary inputs for the leader and ith follower, respectively. Note that v_0 and v_i will not be used in the robust control process. However, they are critical in the learning process to help construct the optimal policies u_0 and u_i.

When $g_j(x)$, $j \in \mathcal{P}$, is a real matrix, we have $g_j^+(x) = (g_j^T(x)g_j(x))^{-1}g_j^T(x)$. Hence, based on Assumption 2, we can further obtain

$$g_j^+(x)k_j(x) = \left(g_j^T(x)g_j(x)\right)^{-1}g_j^T(x)k_j(x) = 0. \tag{3}$$

Therefore, by defining $\vartheta_j = [u_j^T, v_j^T]^T$ and $\mathcal{G}_j(x) = [g_j(x), k_j(x)]$, the auxiliary plant (2) becomes

$$\dot{x} = f(x) + \mathcal{G}_0(x)\vartheta_0 + \sum_{i=1}^{N}\mathcal{G}_i(x)\vartheta_i \tag{4}$$

which is a multiplayer Stackelberg–Nash game. Comparing with (1), the transformed system (4) also involves multiple players in hierarchy: the player 0 as the leader who takes the action first and the other players $i \in \mathcal{F}$ as the followers who make the decisions later by considering the leader's decisions. The control input function $\mathcal{G}_j(x)$, $j \in \mathcal{P}$, is assumed upper bounded as $||\mathcal{G}_j(x)|| \leq \mathcal{G}_M$. Since $\mathcal{G}_j(x) = [g_j(x), k_j(x)]$, we have $||g_j(x)|| \leq \mathcal{G}_M$. In this article, we demonstrate that the solution obtained from system (4) can asymptotically stabilize the original uncertain system (1) if an appropriate cost function is established for each player (see Section III-C, Theorem 1).

Assume the augment system (4) is controllable. Construct the cost function associated to each player $j \in \mathcal{P}$ as

$$J_j(x, \vartheta_j, \vartheta_{-j}) = \int_0^\infty \Big\{\Psi_j^2(x(\tau)) + \mathcal{R}_j\big(x(\tau), \vartheta_j(\tau), \vartheta_{-j}(\tau)\big)\Big\}d\tau \tag{5}$$

where $\Psi_j(x)$ is the design parameter, $R_j(x, \vartheta_j, \vartheta_{-j})$ is the utility function for the jth player, and $\vartheta_{-j} = [u_{-j}^T, v_{-j}^T]^T$ with $u_{-j} = \{u_\xi | \xi \in \mathcal{P}, \xi \neq j\}$ and $v_{-j} = \{v_\xi | \xi \in \mathcal{P}, \xi \neq j\}$.

Define the utility function for the leader as

$$\mathcal{R}_0(x, \vartheta_0, \vartheta_{-0}) = x^T Q_0 x + \left\|\vartheta_0 + \sum_{i=1}^{N}\alpha_i\vartheta_i\right\|_{\mathcal{M}_0}^2 \tag{6}$$

where $\alpha_i \in \mathbb{R}^{(m_0+l_0)\times(m_i+l_i)}$ denotes the coupling coefficient of follower i to the leader and $\alpha_i = \text{diag}\{\alpha_{i1}, \alpha_{i2}\}$ with $\alpha_{i1} \in \mathbb{R}^{m_0\times m_i}$ and $\alpha_{i2} \in \mathbb{R}^{l_0\times l_i}$, $\alpha_{i1} > 0$ and $\alpha_{i2} > 0$. Besides, $Q_0 \in \mathbb{R}^{n\times n}$ is the symmetric positive-definite matrix, and $\mathcal{M}_0 \in \mathbb{R}^{(m_0+l_0)\times(m_0+l_0)}$ is the diagonal positive-definite matrix as $\mathcal{M}_0 = \text{diag}\{I_{m_0}, \rho_0\}$ with $\rho_0 = \text{diag}\{\rho_{01}, \rho_{02}, \ldots, \rho_{0l_0}\}$, $\rho_{0y} > 0$, $y = 1, 2, \ldots, l_0$.

Authorized licensed use limited to: Florida Atlantic University. Downloaded on February 18,2026 at 16:15:36 UTC from IEEE Xplore. Restrictions apply.

For the follower $i \in \mathcal{F}$, define

$$\mathcal{R}_i\ \ (x, \vartheta_i, \vartheta_{-i}) = x^T Q_i x + \left\| \vartheta_i + \beta_i \vartheta_0 \right\|^2_{\mathcal{M}_i} \tag{7}$$

where $\beta_i \in \mathbb{R}^{(m_i+l_i)\times(m_0+l_0)}$ represents the coupling coefficient of the leader to follower i, which is defined as $\beta_i = \text{diag}\{\beta_{i1}, \beta_{i2}\}$ with $\beta_{i1} \in \mathbb{R}^{m_i \times m_0}$ and $\beta_{i2} \in \mathbb{R}^{l_i \times l_0}$. We have $\beta_{i1} > 0$ and $\beta_{i2} > 0$. Furthermore, $Q_i \in \mathbb{R}^{n\times n}$ is the positive-definite symmetric matrix, and $\mathcal{M}_i \in \mathbb{R}^{(m_i+l_i)\times(m_i+l_i)}$ is the diagonal positive-definite matrix as $\mathcal{M}_i = \text{diag}\{I_{m_i}, \rho_i\}$ with $\rho_i = \text{diag}\{\rho_{i1}, \rho_{i2}, \ldots, \rho_{il_i}\}$, $\rho_{iy_1} > 0$, $y_1 = 1, 2, \ldots, l_i$.

Definition 1 (Stackelberg–Nash Equilibrium): If there exists a mapping $\mathcal{D}_i : \mathcal{U}_0 \to \mathcal{U}_i, i \in \mathcal{F}$, such that for a given control policy of the leader $\vartheta_0 \in \mathcal{U}_0$, $\vartheta_i' = \mathcal{D}_i(\vartheta_0)$ is the optimal control policy for the ith follower. If for any given $\vartheta_0 \in \mathcal{U}_0$

$$J_i(x, \mathcal{D}_i(\vartheta_0), \mathcal{D}_{-i}(\vartheta_0)) \le J_i(x, \vartheta_i, \mathcal{D}_{-i}(\vartheta_0)),\ \ \vartheta_i \in \mathcal{U}_i \tag{8}$$

and if there exists ϑ_0' for the leader that

$$J_0\big(x, \vartheta_0', \mathcal{D}_{-0}(\vartheta_0')\big) \le J_0(x, \vartheta_0, \mathcal{D}_{-0}(\vartheta_0)),\ \ \vartheta_0 \in \mathcal{U}_0 \tag{9}$$

where $\mathcal{D}_{-i}(\cdot) = \{\mathcal{D}_\xi(\cdot) | \xi \in \mathcal{F}, \xi \ne i\}$ and $\mathcal{D}_{-0} = \{\mathcal{D}_\xi(\cdot) | \xi \in \mathcal{F}\}$, then the set $\{\vartheta_0', \vartheta_1', \ldots, \vartheta_N'\} \in \mathcal{U}_0 \times \mathcal{U}_1 \times \cdots \times \mathcal{U}_N$ is considered to constitute the Stackelberg–Nash equilibrium.

Note that condition (8) of Definition 1 characterizes that the follower i responds to the leader's decision ϑ_0 and engages in a strategic interaction with other followers $\mathcal{D}_{-i}(\vartheta_0)$ to reach a Nash equilibrium. Moreover, condition (9) denotes the Stackelberg equilibrium, in which the leader leverages on its hierarchical position to strategically anticipate how followers will respond. This foresight guides the leader on determining the policy ϑ_0' that can minimize its cost function J_0. Therefore, the dynamics of this hierarchical system reveal a continuous interplay between leader and followers, where the Stackelberg equilibrium for the leader and the Nash equilibrium for the followers are interdependent.

III. Event-Triggered RL Approach Design

A. Coupled HJ Equation

Design the performance index for each player $j \in \mathcal{P}$ as

$$V_j(x) = \int_t^\infty \Big\{ \Psi_j^2(x(\tau)) + \mathcal{R}_j\big(x(\tau), \vartheta_j(\tau), \vartheta_{-j}(\tau)\big) \Big\} d\tau \tag{10}$$

where $\mathcal{R}_j(x, \vartheta_j, \vartheta_{-j})$ is defined in (6) for leader and in (7) for the ith follower. The Hamiltonian with respect to $V_j(x)$ and ϑ_j can be provided as

$$H_j\big(x, \nabla V_j, \vartheta_j, \vartheta_{-j}\big) = \Psi_j^2(x) + \mathcal{R}_j\big(x, \vartheta_j, \vartheta_{-j}\big) + \nabla V_j^T(x)\Big(f(x) + \mathcal{G}_0(x)\vartheta_0 + \sum_{\xi=1}^N \mathcal{G}_\xi(x)\vartheta_\xi\Big) \tag{11}$$

where $\nabla V_j(x) = ([\partial V_j(x)]/\partial x)$. Hence, we have the optimal performance index as

$$V_j^*(x) = \min_{\vartheta_j \in \mathcal{U}_j} V_j(x), \quad j \in \mathcal{P} \tag{12}$$

which satisfies the coupled HJ equation

$$H_j\Big(x, \nabla V_j^*, \vartheta_j^*, \vartheta_{-j}^*\Big) = 0, \quad j \in \mathcal{P} \tag{13}$$

with $V_j^*(0) = 0$.

Based on (13), we can further derive the coupled control policies for leader and followers, respectively. Specifically, since the leader has a strategic advantage, which will impact the responses of the followers, given the leader's control policy ϑ_0, the optimal control policy of the ith follower can be determined as

$$\begin{aligned} \vartheta_i^{\vartheta_0} &= \arg\min_{\vartheta_i \in \mathcal{U}_i} H_i\Big(x, \nabla V_i^{\vartheta_0}, \vartheta_i, \vartheta_{-i}\Big) \\ &= -\beta_i \vartheta_0 - \frac{1}{2}\mathcal{M}_i^{-1}\mathcal{G}_i^T(x)\nabla V_i^{\vartheta_0}(x) \end{aligned} \tag{14}$$

where $\nabla V_i^{\vartheta_0}(x) = ([\partial V_i^{\vartheta_0}(x)]/\partial x)$, and $V_i^{\vartheta_0}(x)$ is the performance index of the ith follower given ϑ_0.

Conversely, the optimal control policy of the leader is determined by taking into account the followers' best responses. We have

$$\begin{aligned} \vartheta_0^* &= \arg\min_{\vartheta_0 \in \mathcal{U}_0} H_0\Big(x, \nabla V_0^*, \vartheta_0, \vartheta_{-0}^{\vartheta_0}\Big) \\ &= -K_1\left(\mathcal{G}_0(x) - \sum_{i=1}^N \mathcal{G}_i(x)\beta_i\right)^T \nabla V_0^*(x) \\ &\quad + K_2 \sum_{i=1}^N \alpha_i \mathcal{M}_i^{-1}\mathcal{G}_i(x)\nabla V_i^*(x) \end{aligned} \tag{15}$$

where

$$K_1 = \frac{1}{2}\left(\left(I_{m_0+l_0} - \sum_{i=1}^N \alpha_i\beta_i\right)^T \mathcal{M}_0 \left(I_{m_0+l_0} - \sum_{i=1}^N \alpha_i\beta_i\right)\right)^{-1}$$

$$K_2 = \frac{1}{2}\left(I_{m_0+l_0} - \sum_{i=1}^N \alpha_i\beta_i\right)^{-1}.$$

Considering the definition of ϑ_0, $\mathcal{G}_0(x)$, and $\mathcal{G}_i(x)$, we can rewrite (15) as

$$\begin{cases} u_0^* = -K_{11}\left(g_0(x) - \sum_{i=1}^N g_i(x)\beta_{i1}\right)^T \nabla V_0^*(x) \\ \qquad\quad + K_{21}\sum_{i=1}^N \alpha_{i1} g_i(x)\nabla V_i^*(x) \\ v_0^* = -K_{12}\left(k_0(x) - \sum_{i=1}^N k_i(x)\beta_{i2}\right)^T \nabla V_0^*(x) \\ \qquad\quad + K_{22}\sum_{i=1}^N \alpha_{i2}\rho_i^{-1} g_i(x)\nabla V_i^*(x) \end{cases} \tag{16}$$

where

$$K_{11} = \frac{1}{2}\left(\left(I_{m_0} - \sum_{i=1}^N \alpha_{i1}\beta_{i1}\right)^T \left(I_{m_0} - \sum_{i=1}^N \alpha_{i1}\beta_{i1}\right)\right)^{-1}$$

$$K_{12} = \frac{1}{2}\left(\left(I_{l_0} - \sum_{i=1}^N \alpha_{i2}\beta_{i2}\right)^T \rho_i \left(I_{l_0} - \sum_{i=1}^N \alpha_{i2}\beta_{i2}\right)\right)^{-1}$$

$$K_{21} = \frac{1}{2}\left(I_{m_0} - \sum_{i=1}^N \alpha_{i1}\beta_{i1}\right)^{-1}$$

$$K_{22} = \frac{1}{2}\left(I_{l_0} - \sum_{i=1}^N \alpha_{i2}\beta_{i2}\right)^{-1}.$$

It is easy to select appropriate coefficients such that

Authorized licensed use limited to: Florida Atlantic University. Downloaded on February 18,2026 at 16:15:36 UTC from IEEE Xplore. Restrictions apply.

$\sum_{i=1}^{N}\alpha_{i1}\beta_{i1}\neq I_{m_0}$ and $\sum_{i=1}^{N}\alpha_{i2}\beta_{i2}\neq I_{l_0}$.

Substituting (15) into (14), we have the optimal response of the ith follower given the optimal leader's policy ϑ_0^* as

$$\vartheta_i^* = -\beta_i\vartheta_0^* - \frac{1}{2}\mathcal{M}_i^{-1}\mathcal{G}_i^T(x)\nabla V_i^*(x) \tag{17}$$

i.e.,

$$\begin{cases} u_i^* = -\beta_{i1}u_0^* - \frac{1}{2}g_i(x)\nabla V_i^*(x) \\ v_i^* = -\beta_{i2}v_0^* - \frac{1}{2}\rho_i^{-1}k_j(x)\nabla V_i^*(x). \end{cases} \tag{18}$$

It is noteworthy that the control policy of the leader (15) involves the responses of all followers, while each follower's control policy (17) also includes an additional term related to the leader. This coupling property between the leader and followers introduces a complex interdependency within the multiplayer game, which makes the overall design challenging.

B. Design Event-Triggered Control Within the Framework of Stackelberg–Nash Game

To alleviate the computational burden, we design an event-triggered control scheme for the augment system (4) within the framework of Stackelberg–Nash game.

Let $\{t_s\}_{s=0}^{\infty}$ be the sequence of triggering instants, where t_s represents the sth triggering instant and $t_s < t_{s+1}$, $s\in\mathbb{N}$. At the triggering instant t_s, we have the sampled system state as $\bar{x}_s = x(t_s)$, $s\in\mathbb{N}$. Hence, the event-triggered control policy $\vartheta_j(\bar{x}_s)$ for player $j\in\mathcal{P}$, designed based on the sampled state $\bar{x}_s$, is the intermittent feedback. There exists a gap between the sampled and current state, which is given as

$$e_s = x(t) - \bar{x}_s, \quad t\in[t_s, t_{s+1}). \tag{19}$$

This means when the event is triggered ($t = t_s$), we have $e_s = 0$. The control policy is only updated at the triggered instants and kept as the same with a zero-order hold (ZOH) until the event is triggered again. In this way, the control signals in the sequence $\vartheta_j(\bar{x}_s)$ can be converted into a continuous-time signal $\pi_j(\bar{x}_s, t)$ as

$$\pi_j(\bar{x}_s, t) = \vartheta_j(\bar{x}_s), \quad t\in[t_s, t_{s+1}). \tag{20}$$

Applying this event-triggered control policy (20) on the transformed augment system (4), we have

$$\dot{x} = f(x) + \mathcal{G}_0(x)\pi_0(\bar{x}_s, t) + \sum_{i=1}^{N}\mathcal{G}_i(x)\pi_i(\bar{x}_s, t). \tag{21}$$

This event-triggered control design is within the framework of Stackelberg–Nash game. Specifically, all the players are triggered at the same time instant $t = t_s$, $s\in\mathbb{N}$. The hierarchical structure persists during this time instant, that is, the leader $\pi_0(\bar{x}_s, t) = \vartheta_0(\bar{x}_s)$ acts first, and the followers $\pi_i(\bar{x}_s, t) = \vartheta_i(\bar{x}_s)$ respond to the leader's decision subsequently. In this way, we further convert this robust hierarchical control problem into an event-triggered optimization design of Stackelberg–Nash game.

Considering (15) and (17), the optimal event-triggered control policy for the leader is given as

$$\pi_0^*(\bar{x}_s, t) = \vartheta_0^*(\bar{x}_s) = -K_1\left(\mathcal{G}_0(\bar{x}_s) - \sum_{i=1}^{N}\mathcal{G}_i(\bar{x}_s)\beta_i\right)^T \cdot\nabla V_0^*(\bar{x}_s) + K_2\sum_{i=1}^{N}\alpha_i\mathcal{M}_i^{-1}\mathcal{G}_i(\bar{x}_s)\nabla V_i^*(\bar{x}_s) \tag{22}$$

and the optimal event-triggered control policy for the ith follower is provided as

$$\pi_i^*(\bar{x}_s, t) = \vartheta_i^*(\bar{x}_s) = -\beta_i\vartheta_0^*(\bar{x}_s) - \frac{1}{2}\mathcal{M}_i^{-1}\mathcal{G}_i^T(\bar{x}_s)\nabla V_i^*(\bar{x}_s) \tag{23}$$

where $\nabla V_0^*(\bar{x}_s) = ([\partial V_0^*(x)]/\partial x)|_{x=\bar{x}_s}$ and $\nabla V_i^*(\bar{x}_s) = ([\partial V_i^*(x)]/\partial x)|_{x=\bar{x}_s}$.

Based on the definition of ϑ_0^* and ϑ_i^*, we can further derive that

$$\pi_{0,u}^*(\bar{x}_s, t) = u_0^*(\bar{x}_s) = -K_{11}\left(g_0(\bar{x}_s) - \sum_{i=1}^{N}g_i(\bar{x}_s)\beta_{i1}\right)^T \cdot\nabla V_0^*(\bar{x}_s) + K_{21}\sum_{i=1}^{N}\alpha_{i1}g_i(\bar{x}_s)\nabla V_i^*(\bar{x}_s) \tag{24}$$

$$\pi_{i,u}^*(\bar{x}_s, t) = u_i^*(\bar{x}_s) = -\beta_{i1}u_0^*(\bar{x}_s) - \frac{1}{2}g_i(\bar{x}_s)\nabla V_i^*(\bar{x}_s). \tag{25}$$

Substituting (22) and (23) into (13), we obtain the coupled event-triggered HJ equation for each player $j\in\mathcal{P}$ as

$$H_j\left(x, \nabla V_j^*, \vartheta_j^*(\bar{x}_s), \vartheta_{-j}^*(\bar{x}_s)\right) = \mathcal{R}_j\left(x, \vartheta_j^*(\bar{x}_s), \vartheta_{-j}^*(\bar{x}_s)\right) + \Psi_j^2(x) + \left(\nabla V_j^*(x)\right)^T\left(f(x) + \sum_{\xi=0}^{N}\mathcal{G}_\xi(x)\vartheta_\xi^*(\bar{x}_s)\right). \tag{26}$$

C. Stability Preservation of the Designed Transformation

Assumption 3: The control policy ϑ_j^* for each player satisfies the Lipschitz condition. Therefore, we can find a Lipschitz constant $\mathcal{L}_{\vartheta_j} > 0$ such that

$$||\vartheta_j^*(x) - \vartheta_j^*(\bar{x}_s)|| \leq \mathcal{L}_{\vartheta_j}||x(t) - \bar{x}_s|| = \mathcal{L}_{\vartheta_j}||e_s||. \tag{27}$$

Note that due to the fact $\vartheta_j = [u_j^T, v_j^T]^T$ and the definition of norm $||\cdot||$, we can also derive

$$||u_j^*(x) - u_j^*(\bar{x}_s)|| \leq \mathcal{L}_{\vartheta_j}||e_s||. \tag{28}$$

Define $\mathcal{L}_\vartheta = \sum_{j=0}^{N}\mathcal{L}_{\vartheta_j}$. Lipschitz continuity ensures predictable and bounded control behavior.

Theorem 1: Consider the optimal event-triggered control policies $\vartheta_0^*(\bar{x}_s)$ in (22) and $\vartheta_i^*(\bar{x}_s), i\in\mathcal{F}$, in (23). Assume the validity of Assumption 1–3. Let $V_j^*(x), j\in\mathcal{P}$, be the performance index, with the parameter $\Psi_j^2(x)$ defined as

$$\Psi_j^2(x) = \frac{3}{4}\nabla V_j^{*T}(x)\nabla V_j^*(x) + \mathscr{D}_d^2(x) + \mathscr{B}_v^2(x) \tag{29}$$

Authorized licensed use limited to: Florida Atlantic University. Downloaded on February 18,2026 at 16:15:36 UTC from IEEE Xplore. Restrictions apply.

where $\mathscr{B}_v^2(x)$ is an upper bound given as $\mathscr{B}_v(x) \geq \| \sum_{j=0}^{N} k_j(x) v_j^*(x) \|$. If the triggering condition is designed as

$$||e_s||^2 \leq \frac{\sum_{j=0}^{N} \left(1 - a_j^2\right) \lambda_{\min}(Q_j) ||x||^2}{(N+1) \mathcal{G}_M^2 \mathscr{L}_\vartheta^2} \tag{30}$$

where $a_j \in (0, 1)$ and $\lambda_{\min}(Q_j)$ denotes the minimum eigenvalue of Q_j, then the policies $\vartheta_0^*(\bar{x}_s)$ and $\vartheta_i^*(\bar{x}_s)$ ensure the asymptotic stability of the hierarchical mismatched uncertain system (1).

Proof: Choose the Lyapunov function candidate as

$$L(x) = \sum_{j=0}^{N} V_j^*(x). \tag{31}$$

Taking the time derivative of $L(x)$ in (31) along the trajectory of system (1) with the event-triggered control policies (22) for the leader and (23) for the followers, we obtain

$$\dot{L}(x) = \sum_{j=0}^{N} \left\{ \nabla V_j^{*T}(x) \left(f(x) + \sum_{\xi=0}^{N} g_\xi(x) u_\xi^*(\bar{x}_s) + \sum_{\xi=0}^{N} k_\xi(x) d_\xi(x) \right) \right\}. \tag{32}$$

According to (13), we have

$$\nabla V_j^{*T}(x) f(x) = -\Psi_j^2(x) - \mathcal{R}_j\left(x, \vartheta_j^*, \vartheta_{-j}^*\right) - \nabla V_j^{*T}(x) \sum_{\xi=0}^{N} \mathcal{G}_\xi(x) \vartheta_\xi^*(x). \tag{33}$$

Substituting (33) into (32), it follows:

$$\dot{L}(x) = \sum_{j=0}^{N} \left\{ -\Psi_j^2(x) - \mathcal{R}_j\left(x, \vartheta_j^*, \vartheta_{-j}^*\right) - \nabla V_j^{*T}(x) \cdot \sum_{\xi=0}^{N} \mathcal{G}_\xi(x) \vartheta_\xi^*(x) + \nabla V_j^{*T}(x) \sum_{\xi=0}^{N} g_\xi(x) u_\xi^*(\bar{x}_s) + \nabla V_j^{*T}(x) \sum_{\xi=0}^{N} k_\xi(x) d_\xi(x) \right\}. \tag{34}$$

Considering the definition of $\vartheta_j^* = [u_j^{*T}, v_j^{*T}]^T$ and $\mathcal{G}_j(x) = [g_j(x), k_j(x)]$, we can further rewrite (34) as

$$\begin{aligned} \dot{L}(x) &= \sum_{j=0}^{N} \left\{ -\Psi_j^2(x) - \mathcal{R}_j\left(x, \vartheta_j^*, \vartheta_{-j}^*\right) + \nabla V_j^{*T}(x) \cdot \left(\sum_{\xi=0}^{N} g_\xi(x) \left(u_\xi^*(\bar{x}_s) - u_\xi^*(x)\right) + \sum_{\xi=0}^{N} k_\xi(x) \left(d_\xi(x) - v_\xi^*(x)\right) \right) \right\} \\ &\leq \sum_{j=0}^{N} \left\{ -\mathcal{R}_j\left(x, \vartheta_j^*, \vartheta_{-j}^*\right) - \mathscr{D}_d^2(x) - \mathscr{B}_v^2(x) + \left\| \sum_{\xi=0}^{N} g_\xi(x) \left(u_\xi^*(\bar{x}_s) - u_\xi^*(x)\right) \right\|^2 + \left\| \sum_{\xi=0}^{N} k_\xi(x) d_\xi(x) \right\|^2 + \left\| \sum_{\xi=0}^{N} k_\xi(x) v_\xi^*(x) \right\|^2 \right\}. \end{aligned} \tag{35}$$

Based on the bounded conditions, it follows:

$$\dot{L}(x) \leq \sum_{j=0}^{N} \left\{ -a_j^2 \lambda_{\min}(Q_j) ||x||^2 - \left[\left(1 - a_j^2\right) \lambda_{\min}(Q_j) ||x||^2 - (N+1) \mathcal{G}_M^2 \mathscr{L}_\vartheta^2 ||e_s||^2 \right] \right\}. \tag{36}$$

If the triggering condition (30) holds, we obtain $\dot{L}(x) \leq \sum_{j=0}^{N} -a_j^2 \lambda_{\min}(Q_j) ||x||^2 < 0$ for $\forall x \neq 0$. Therefore, the designed event-triggered control scheme is guaranteed to asymptotically stabilize the hierarchical system with mismatched uncertainties (1). This concludes the proof. ■

Theorem 1 demonstrates that the event-triggered optimal control solution for the transformed Stackelberg–Nash game guarantees the asymptotic stabilization of the original hierarchical robust control system.

Based on (30), we can derive the triggering instant t_s. Subsequently, the minimal intersampling interval can be calculated as $T_{s,\min} = (t_{s+1} - t_s)_{\min}$. For the continuous-time systems with an event-triggered controller design, it is crucial to ensure $T_{s,\min} \neq 0$, i.e., free from Zeno behavior. Therefore, consider the event-triggered mechanism, at the sth triggering instant $e_s = 0$, the time of $(||e_s||/||x||)$ growing from 0 to $P_{\text{threshold}}$ provides a lower bound for the minimal intersampling interval. Here, we can obtain $P_{\text{threshold}} = (1/[\mathscr{L}_\vartheta ||\mathcal{G}_M||]) \sqrt{([\sum_{j=0}^{N} (1 - a_j^2) \lambda_{\min}(Q_j)]/(N+1))} > 0$. This means the designed method can guarantee $T_{s,\min} > 0$ and achieve Zeno-free behavior.

Theorem 2: Let $V_j^*(x)$ be the optimal performance index with the parameter $\Psi_j^2(x)$ defined as (29). If the triggering condition is defined as (30) and Assumption 1–3 hold, then the designed event-triggered control policy set $\{\vartheta_0^*(\bar{x}_s), \vartheta_1^*(\bar{x}_s), \ldots, \vartheta_N^*(\bar{x}_s)\}$ can achieve Stackelberg–Nash equilibrium.

Proof: Based on Theorem 1, we have $x \to 0$ when $t \to \infty$ with the designed event-triggered control policies. The optimal performance index for each player $j \in \mathcal{P}$ satisfies $V_j^*(x(\infty)) = V_j^*(0) = 0$. Set a new parameter $\mathcal{Z}_j(x) = V_j^*(x)$. Therefore, the cost function (5) under the event-triggered control policy $\vartheta_j(\bar{x}_s)$ can be rewritten as

$$J_j\left(x, \vartheta_j(\bar{x}_s), \vartheta_{-j}(\bar{x}_s)\right) = \int_0^\infty \left\{ \mathcal{R}_j\left(x(\tau), \vartheta_j(\bar{x}_s), \vartheta_{-j}(\bar{x}_s)\right) + \Psi_j^2(x(\tau)) \right\} d\tau + \mathcal{Z}_j(x(0)) + \int_0^\infty \dot{\mathcal{Z}}_j(x(\tau)) d\tau \tag{37}$$

where $\dot{\mathcal{Z}}_j(x) = \nabla \mathcal{Z}_j^T(x)(f(x) + \sum_{\xi=0}^{N} \mathcal{G}_\xi(x) \vartheta_\xi(\bar{x}_s))$ and $\nabla \mathcal{Z}_j(x) = ([\partial \mathcal{Z}_j(x)]/\partial x)$. Considering (26), we can obtain that at the triggering instant

$$J_j\left(x, \vartheta_j(\bar{x}_s), \vartheta_{-j}(\bar{x}_s)\right) = \int_0^\infty \left\{ \mathcal{R}_j\left(x(\tau), \vartheta_j(\bar{x}_s), \vartheta_{-j}(\bar{x}_s)\right) - \mathcal{R}_j\left(x(\tau), \vartheta_j^*(\bar{x}_s), \vartheta_{-j}^*(\bar{x}_s)\right) + \nabla \mathcal{Z}_j^T(\bar{x}_s) \sum_{\xi=0}^{N} \mathcal{G}_\xi(\bar{x}_s) \cdot \left(\vartheta_\xi(\bar{x}_s) - \vartheta_\xi^*(\bar{x}_s)\right) \right\} d\tau + \mathcal{Z}_j(x(0)). \tag{38}$$

Authorized licensed use limited to: Florida Atlantic University. Downloaded on February 18,2026 at 16:15:36 UTC from IEEE Xplore. Restrictions apply.

For the leader, assume that all the followers take the optimal control policies $\vartheta_{-0}^{\vartheta_0}(\bar{x}_s) = \{\vartheta_\xi^{\vartheta_0}(\bar{x}_s)|\xi \in \mathcal{F}\}$ given $\vartheta_0(\bar{x}_s)$. Therefore, we have

$$\begin{aligned} &J_0\Big(x, \vartheta_0(\bar{x}_s), \vartheta_{-0}^{\vartheta_0}(\bar{x}_s)\Big) \\ &= \mathcal{Z}_0(x(0)) + \int_0^\infty \Bigg\{ \Bigg\| \vartheta_0(\bar{x}_s) + \sum_{\xi=1}^N \alpha_\xi \vartheta_\xi^{\vartheta_0}(\bar{x}_s) \Bigg\|_{\mathcal{M}_0}^2 \\ &- \Bigg\| \vartheta_0^*(\bar{x}_s) + \sum_{\xi=1}^N \alpha_\xi \vartheta_\xi^*(\bar{x}_s) \Bigg\|_{\mathcal{M}_0}^2 + \nabla \mathcal{Z}_0^T(\bar{x}_s) \mathcal{G}_0(\bar{x}_s)(\vartheta_0(\bar{x}_s) \\ &- \vartheta_0^*(\bar{x}_s)\Big) + \nabla \mathcal{Z}_0^T(\bar{x}_s) \sum_{\xi=1}^N \mathcal{G}_\xi(\bar{x}_s)\Big(\vartheta_\xi^{\vartheta_0}(\bar{x}_s) - \vartheta_\xi^*(\bar{x}_s)\Big)\Bigg\} d\tau \\ &= \mathcal{Z}_0(x(0)) + \int_0^\infty \Big\{ H_0\Big(x, \nabla \mathcal{Z}_0(\bar{x}_s), \vartheta_0(\bar{x}_s), \vartheta_{-0}^{\vartheta_0}(\bar{x}_s)\Big) \\ &- H_0\Big(x, \nabla \mathcal{Z}_0(\bar{x}_s), \vartheta_0^*(\bar{x}_s), \vartheta_{-0}^*(\bar{x}_s)\Big)\Big\}. \end{aligned} \tag{39}$$

At the equilibrium point $\vartheta_0(\bar{x}_s) = \vartheta_0^*(\bar{x}_s)$, the integral term in (39) becomes zero. It follows:

$$J_0\Big(x, \vartheta_0^*(\bar{x}_s), \vartheta_{-0}^*(\bar{x}_s)\Big) = \mathcal{Z}_0(x(0)). \tag{40}$$

Since the goal of $\vartheta_0(\bar{x}_s)$ is to minimize H_0, we have $H_0(x, \nabla \mathcal{Z}_0(\bar{x}_s), \vartheta_0(\bar{x}_s), \vartheta_{-0}^{\vartheta_0}(\bar{x}_s)) \geq H_0(x, \nabla \mathcal{Z}_0(\bar{x}_s), \vartheta_0^*(\bar{x}_s), \vartheta_{-0}^*(\bar{x}_s))$. Combining this with (40), we obtain

$$J_0\Big(x, \vartheta_0(\bar{x}_s), \vartheta_{-0}^{\vartheta_0}(\bar{x}_s)\Big) \geq J_0\Big(x, \vartheta_0^*(\bar{x}_s), \vartheta_{-0}^*(\bar{x}_s)\Big). \tag{41}$$

Now, we consider the situation for the ith follower. Assume that the other followers take the optimal policies given $\vartheta_0^*(\bar{x}_s)$ as $\vartheta_{-i}^*(\bar{x}_s)) = \{\vartheta_\xi^{\vartheta_0^*}(\bar{x}_s) = \vartheta_\xi^*(\bar{x}_s)|\xi \in \mathcal{F}, \xi \neq i\}$. Based on (38), we obtain

$$\begin{aligned} J_i\Big(x, \vartheta_i(\bar{x}_s), \vartheta_{-i}^*(\bar{x}_s)\Big) &= \mathcal{Z}_i(x(0)) + \int_0^\infty \Bigg\{ \Big\| \vartheta_i(\bar{x}_s) \\ &+ \beta_i \vartheta_0^*(\bar{x}_s) \Big\|_{\mathcal{M}_i}^2 - \Big\| \vartheta_i^*(\bar{x}_s) + \vartheta_0^*(\bar{x}_s) \Big\|_{\mathcal{M}_i}^2 \\ &+ \nabla \mathcal{Z}_i^T(\bar{x}_s) \mathcal{G}_i(\bar{x}_s)\Big(\vartheta_i(\bar{x}_s) - \vartheta_i^*(\bar{x}_s)\Big)\Bigg\} d\tau. \end{aligned} \tag{42}$$

According to (23), it holds

$$\nabla \mathcal{Z}_i^T(\bar{x}_s) \mathcal{G}_i^T(\bar{x}_s) = -2\Big(\vartheta_i^*(\bar{x}_s) + \beta_i \vartheta_0^*(\bar{x}_s)\Big)^T \mathcal{M}_i. \tag{43}$$

Substituting (43) into (42) and completing the square, it follows

$$\begin{aligned} J_i\Big(x, \vartheta_i(\bar{x}_s), \vartheta_{-i}^*(\bar{x}_s)\Big) &= \mathcal{Z}_i(x(0)) \\ &+ \int_0^\infty \Big\{\Big(\vartheta_i(\bar{x}_s) - \vartheta_i^*(\bar{x}_s)\Big)\mathcal{M}_i\Big(\vartheta_i(\bar{x}_s) - \vartheta_i^*(\bar{x}_s)\Big)\Big\} d\tau \end{aligned} \tag{44}$$

It is clear that when $\vartheta_i(\bar{x}_s) = \vartheta_i^*(\bar{x}_s)$, we have $J_i(x, \vartheta_i^*(\bar{x}_s), \vartheta_{-i}^*(\bar{x}_s)) = \mathcal{Z}_i(x(0))$. Hence, from (44), one has $J_i(x, \vartheta_i(\bar{x}_s), \vartheta_{-i}^*(\bar{x}_s)) \geq \mathcal{Z}_i(x(0))$, or

$$J_i\Big(x, \vartheta_i(\bar{x}_s), \vartheta_{-i}^*(\bar{x}_s)\Big) \geq J_i\Big(x, \vartheta_i^*(\bar{x}_s), \vartheta_{-i}^*(\bar{x}_s)\Big). \tag{45}$$

Combine (45) and (41), and compare them with the conditions (8) and (9) in Definition (1), respectively. We obtain that the designed event-triggered control policy set $\{\vartheta_0^*(\bar{x}_s), \vartheta_1^*(\bar{x}_s), \ldots, \vartheta_N^*(\bar{x}_s)\}$ can achieve Stackelberg–Nash equilibrium. ■

Remark 1: Consider the fact that $\vartheta_0(\bar{x}_s) = [u_0^T(\bar{x}_s), v_0^T(\bar{x}_s)]^T$ and $\vartheta_i(\bar{x}_s) = [u_i^T(\bar{x}_s), v_i^T(\bar{x}_s)]^T$. Based on Theorem 2 and Definition 2, we can easily obtain that the derived event-triggered control policy set $\{u_0^*(\bar{x}_s), u_1^*, \ldots, u_N^*(\bar{x}_s)\}$ also achieves Stackelberg–Nash equilibrium.

D. Neural-RL-Based Approach With Stability Guarantee

Since the developed event-triggered control policies (22) and (23) require the knowledge of $V_j^*(\bar{x}_s)$, $j \in \mathcal{P}$ which is difficult to solve directly, this article develops a neural-RL-based approach to automatically learn the optimal strategies for both the leader and the followers.

Establish a critic network for each player $j \in \mathcal{P}$ to approximate the optimal performance index as

$$V_j^*(x) = \omega_{cj}^{*T} \phi_{cj}(x) + \epsilon_{cj}(x) \tag{46}$$

where ω_{cj}^* are the ideal critic network weights, ϕ_{cj} is the activation function, and ϵ_{cj} is the function reconstruction error. The partial derivative of $V_j^*(x)$ with respect to x can be provided as

$$\nabla V_j^*(x) = \nabla \phi_{cj}^T(x) \omega_{cj}^* + \nabla \epsilon_{cj}(x) \tag{47}$$

where $\nabla \phi_{cj}(x) = ([\partial \phi_{cj}(x)]/\partial x)$ and $\nabla \epsilon_{cj}(x) = ([\partial \epsilon_{cj}(x)]/\partial x)$. Substituting (47) into (26), it follows:

$$\Psi_j^2(x) + \mathcal{R}_j\Big(x, \vartheta_j(\bar{x}_s), \vartheta_{-j}(\bar{x}_s)\Big) = -\omega_{cj}^{*T} \lambda_j + \Delta_{H_j} \tag{48}$$

where $\Delta_{H_j} = -\nabla \epsilon_j(x)(f(x) + \sum_{\xi=0}^N \mathcal{G}_\xi(x)\vartheta_\xi(\bar{x}_s))$ is the residual error due to the function approximation and $\lambda_j = \nabla \phi_{cj}(x)(f(x) + \sum_{\xi=0}^N \mathcal{G}_\xi(x)\vartheta_\xi(\bar{x}_s))$.

Assumption 4: For every $x \in \mathbb{R}^n$, $\nabla \phi_{cj}(x)$ and $\nabla \epsilon_{cj}(x)$ are bounded as $||\nabla \phi_{cj}(x)|| \leq \mathscr{B}_{\phi_{cj}}$ and $||\epsilon_{cj}(x)|| \leq \mathscr{B}_{\epsilon_{cj}}$, respectively, with $\mathscr{B}_{\phi_{cj}}$ and $\mathscr{B}_{\epsilon_{cj}}$ being the positive constants, $j \in \mathcal{P}$. In addition, for every $x \in \mathbb{R}^n$, Δ_{H_j} is bounded as $||\Delta_{H_j}|| \leq \mathscr{B}_{\Delta_j}$, with $\mathscr{B}_{\Delta_j}$ being the positive constant, $j \in \mathcal{P}$.

Substituting (47) into (22) and (23), we derive the optimal control policies for leader and ith follower, respectively. However, achieving ω_{cj}^* is often challenging or impractical, which makes the implementation difficult. Therefore, we consider the current estimated value $\hat{\omega}_{cj}$ instead and obtain the approximate performance index as

$$\hat{V}_j(x) = \hat{\omega}_{cj}^T \phi_{cj}(x). \tag{49}$$

The partial derivative of (49) becomes $\nabla \hat{V}_j(x) = \nabla \phi_{cj}^T(x) \hat{\omega}_{cj}$. Note that, at the triggering instant t_s, we have

$$\nabla \hat{V}_j(\bar{x}_s) = \nabla \phi_{cj}^T(\bar{x}_s) \hat{\omega}_{cj} \tag{50}$$

where $\nabla \phi_{cj}(\bar{x}_s) = ([\partial \phi_{cj}(x)]/\partial x)|_{x=\bar{x}_s}$. Replace $\nabla V_0^*(\bar{x}_s)$ in (22) and $\nabla V_i^*(\bar{x}_s)$, $i \in \mathcal{F}$, in (23) with $\nabla \hat{V}_j(\bar{x}_s)$, $j \in \mathcal{P}$, in (50). We attain the estimated event-triggered control policy for the leader as

$$\begin{aligned} \vartheta_0(\bar{x}_s) = &-K_1\Bigg(\mathcal{G}_0(\bar{x}_s) - \sum_{i=1}^N \mathcal{G}_i(\bar{x}_s)\beta_i\Bigg)^T \nabla \phi_{c0}^T(\bar{x}_s) \hat{\omega}_{c0} \\ &+ K_2 \sum_{i=1}^N \alpha_i \mathcal{M}_i^{-1} \mathcal{G}_i(\bar{x}_s) \nabla \phi_{ci}^T(\bar{x}_s) \hat{\omega}_{ci} \end{aligned} \tag{51}$$

Authorized licensed use limited to: Florida Atlantic University. Downloaded on February 18,2026 at 16:15:36 UTC from IEEE Xplore. Restrictions apply.

and for the i th follower as

$$\vartheta_i(\bar{x}_s) = -\beta_i \vartheta_0(\bar{x}_s) - \frac{1}{2}\mathcal{M}_i^{-1}\mathcal{G}_i^T(\bar{x}_s)\nabla\phi_{ci}^T(\bar{x}_s)\hat{\omega}_{ci}. \quad (52)$$

Additionally, with (51) and (52), the estimated form of coupled event-triggered HJ equation can be established as

$$\begin{aligned} &H_j\big(x, \hat{\omega}_{cj}, \vartheta_j(\bar{x}_s), \vartheta_{-j}(\bar{x}_s)\big) \\ &\quad = \Psi_j^2(x) + \mathcal{R}_j\big(x, \vartheta_j(\bar{x}_s), \vartheta_{-j}(\bar{x}_s)\big) + \hat{\omega}_{cj}^T\lambda_j. \end{aligned} \quad (53)$$

Define $e_{cj} = H_j(x, \hat{\omega}_{cj}, \vartheta_j(\bar{x}_s), \vartheta_{-j}(\bar{x}_s))$. Hence, the objective function for the critic network can be designed as

$$E_{cj} = \frac{1}{2}e_{cj}^T e_{cj}. \quad (54)$$

Taking the partial derivative of (54) with respect to $\hat{\omega}_{cj}$, we have the updating rule for the critic network weights as

$$\dot{\hat{\omega}}_{cj} = -\frac{\eta_{cj}\lambda_j e_{cj}}{\left(\lambda_j^T\lambda_j + 1\right)^2} \quad (55)$$

where η_{cj} is the learning rate of the critic network for the jth player, and $(\lambda_j^T\lambda_j + 1)^2$ is the normalization term.

Besides, define the critic network weight estimation error as $\tilde{\omega}_{cj} = \omega_{cj}^* - \hat{\omega}_{cj}$. Based on (53) and (55), we obtain

$$\begin{aligned} \dot{\tilde{\omega}}_{cj} = &\frac{\eta_{cj}\lambda_j}{\left(\lambda_j^T\lambda_j + 1\right)^2} \\ &\cdot \left(\left(\Psi_j^2(x) + \mathcal{R}_j(x, \vartheta_j(\bar{x}_s), \vartheta_{-j}(\bar{x}_s)) + \hat{\omega}_{cj}^T\lambda_j\right). \end{aligned} \quad (56)$$

Substituting (48) into (56), we can further derive the dynamics of $\tilde{\omega}_{cj}$ as

$$\dot{\tilde{\omega}}_{cj} = -\eta_{cj}\gamma_j\gamma_j^T\tilde{\omega}_{cj} + \eta_{cj}\gamma_j\frac{\Delta_{H_j}}{\psi_j} \quad (57)$$

where $\gamma_j = (\lambda_j/[(\lambda_j^T\lambda_j + 1)])$ and $\psi_j = \lambda_j^T\lambda_j + 1$.

The framework diagram of this developed method is provided in Fig. 1, where the leader and followers update their control policies intermittently based on the event-triggered mechanisms. The leader's policy directly influences the followers, while each follower optimizes its strategy by responding to the leader's decision.

Assumption 5: Both ω_{cj}^* and $\hat{\omega}_{cj}$ are upper bounded, i.e., $||\omega_{cj}^*|| \le \mathscr{B}_{\omega_{cj}}$ and $||\hat{\omega}_{cj}|| \le \mathscr{B}_{\omega_{cj}}$, with $\mathscr{B}_{\omega_{cj}}$ being the positive constant, $j \in \mathcal{P}$.

Theorem 3: Consider the augment system (4) with the event-triggered control policy given by (51) for the leader and (52) for the ith follower. Let the critic network weights be updated based on (55). If the triggering condition (30) and Assumption 1–5 hold, then both the closed-loop system (4) and the weight estimation error $\tilde{\omega}_{cj}$ are uniformly ultimate boundedness (UUB).

Proof: Define the Lyapunov function candidate as

$$L_{cl} = \sum_{j=0}^{N} V_j^*(x) + \sum_{j=0}^{N} V_j^*(\bar{x}_s) + \sum_{j=0}^{N}\eta_{cj}^{-1}\mathrm{tr}(\tilde{\omega}_{cj}^T\tilde{\omega}_{cj}) \quad (58)$$

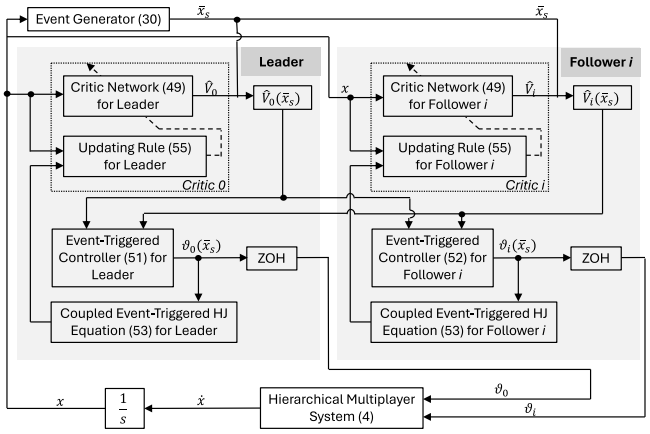

Fig. 1. Framework of the developed event-triggered RL approach. Note: ZOH stands for Zero-Order Hold.

with $L_{clj} = V_j^*(x) + V_j^*(\bar{x}_s) + \eta_{cj}^{-1}\mathrm{tr}(\tilde{\omega}_{cj}^T\tilde{\omega}_{cj})$, $L_{1j} = V_j^*(x)$, $L_{2j} = V_j^*(\bar{x}_s)$, and $L_{3j} = \eta_{cj}^{-1}\mathrm{tr}(\tilde{\omega}_{cj}^T\tilde{\omega}_{cj})$.

This proof is conducted in two parts, i.e., the continuous dynamic phase and the jump dynamic phase. The goal is to demonstrate that both phases of the impulsive closed-loop system are UUB.

We start with the continuous dynamics, $t \in [t_s, t_{s+1})$. By taking the time derivative of (58), we notice $\dot{L}_{2j} = 0$. Therefore, $\dot{L}_{clj} = \dot{L}_{1j} + \dot{L}_{3j}$, which can be further provided as

$$\dot{L}_{clj} = \nabla V_j^{*T}(x)\dot{x}\Big|_{\dot{x}=eq.\ (4)} + \eta_{cj}^{-1}\mathrm{tr}\left(\tilde{\omega}_{cj}^T\dot{\tilde{\omega}}_{cj}\right)\Big|_{\dot{\tilde{\omega}}_{cj}=eq.\ (56)}. \quad (59)$$

Note that $\dot{L}_{cl} = \sum_{j=0}^{N}\dot{L}_{clj}$. Therefore, by demonstrating $\dot{L}_{clj} \le 0$, we can easily obtain $\dot{L}_{cl} \le 0$.

The first term in (59) can be derived as

$$\dot{L}_{1j} = \nabla V_j^{*T}(x)\left(f(x) + \sum_{\xi=0}^{N}\mathcal{G}_\xi(x)\vartheta_\xi(\bar{x}_s)\right). \quad (60)$$

Based on (13), we obtain

$$\begin{aligned} \dot{L}_{1j} = &-\Psi_j^2(x) - R_j\left(x, \vartheta_j^*, \vartheta_{-j}^*\right) \\ &+\nabla V_j^{*T}(x)\sum_{\xi=0}^{N}\mathcal{G}_\xi(x)\left(\vartheta_\xi(\bar{x}_s) - \vartheta_\xi^*\right). \end{aligned} \quad (61)$$

Applying Young's inequality and Cauchy–Schwartz inequality, we can further derive that

$$\begin{aligned} \dot{L}_{1j} \le &-\Psi_j^2(x) - R_j\left(x, \vartheta_j^*, \vartheta_{-j}^*\right) + \frac{1}{2}\left\|\nabla\phi_{cj}^T(x)\omega_{cj}^* + \nabla\epsilon_{cj}(x)\right\|^2 \\ &+ \frac{1}{2}\left\|\sum_{\xi=0}^{N}\mathcal{G}_\xi(x)\left(\vartheta_\xi(\bar{x}_s) - \vartheta_\xi^*\right)\right\|^2 \\ \le &-\Psi_j^2(x) - R_j\left(x, \vartheta_j^*, \vartheta_{-j}^*\right) + \mathscr{B}_{\phi_{cj}}^2\mathscr{B}_{\omega_{cj}}^2 + \mathscr{B}_{\epsilon_{cj}}^2 \\ &+ \frac{1}{2}(N+1)\mathcal{G}_M^2\underbrace{\sum_{\xi=0}^{N}\left\|\vartheta_\xi(\bar{x}_s) - \vartheta_\xi^*\right\|^2}_{\kappa}. \end{aligned} \quad (62)$$

Authorized licensed use limited to: Florida Atlantic University. Downloaded on February 18,2026 at 16:15:36 UTC from IEEE Xplore. Restrictions apply.

The last term in (62) can be rewritten as

$$\kappa = \sum_{\xi=0}^{N}\left\|\left(\vartheta_\xi(\bar{x}_s)-\vartheta_\xi^*(\bar{x}_s)\right)+\left(\vartheta_\xi^*(\bar{x}_s)-\vartheta_\xi^*(x)\right)\right\|^2 \le 2\sum_{\xi=0}^{N}\left\|\vartheta_\xi(\bar{x}_s)-\vartheta_\xi^*(\bar{x}_s)\right\|^2+2\sum_{\xi=0}^{N}\left\|\vartheta_\xi^*(\bar{x}_s)-\vartheta_\xi^*(x)\right\|^2. \quad (63)$$

Based on Assumption 3, we obtain

$$\kappa \le 2\mathscr{L}_\vartheta^2\|e_s\|^2+2\left(\left\|\vartheta_0(\bar{x}_s)-\vartheta_0^*(\bar{x}_s)\right\|^2+\sum_{i=1}^{N}\left\|\vartheta_i(\bar{x}_s)-\vartheta_i^*(\bar{x}_s)\right\|^2\right). \quad (64)$$

Substituting (52) into (64), it follows:

$$\begin{aligned}\kappa &\le 2\mathscr{L}_\vartheta^2\|e_s\|^2+2\Bigg(\left\|\vartheta_0(\bar{x}_s)-\vartheta_0^*(\bar{x}_s)\right\|^2+\sum_{i=1}^{N}\Big\|-\beta_i\big(\vartheta_0(\bar{x}_s)-\vartheta_0^*(\bar{x}_s)\big)-\frac{1}{2}\mathcal{M}_i^{-1}\mathcal{G}_i^T(\bar{x}_s)\big(\nabla\phi_{ci}^T(\bar{x}_s)\tilde{\omega}_{ci}+\epsilon_{ci}\big)\Big\|^2\Bigg)\\ &\le 2\mathscr{L}_\vartheta^2\|e_s\|^2+2\left\|\vartheta_0(\bar{x}_s)-\vartheta_0^*(\bar{x}_s)\right\|^2+4\sum_{i=1}^{N}\Big\|\beta_i\big(\vartheta_0(\bar{x}_s)-\vartheta_0^*(\bar{x}_s)\big)\Big\|^2+4\sum_{i=1}^{N}\Big\|\frac{1}{2}\mathcal{M}_i^{-1}\mathcal{G}_i^T(\bar{x}_s)\big(\nabla\phi_{ci}^T(\bar{x}_s)\tilde{\omega}_{ci}+\epsilon_{ci}\big)\Big\|^2\\ &\le 2\mathscr{L}_\vartheta^2\|e_s\|^2+2\Big(1+2\sum_{i=1}^{N}\|\beta_i\|^2\Big)\underbrace{\left\|\vartheta_0(\bar{x}_s)-\vartheta_0^*(\bar{x}_s)\right\|^2}_{\kappa_{\vartheta_0}}+2\mathcal{G}_M^2\sum_{i=1}^{N}\left\|\mathcal{M}_i^{-1}\right\|^2\big(\mathscr{B}_{\phi_{ci}}^2\|\tilde{\omega}_{ci}\|^2+\mathscr{B}_{\epsilon_{ci}}^2\big).\end{aligned} \quad (65)$$

By taking (51), the term κ_{ϑ_0} in (65) becomes

$$\kappa_{\vartheta_0} = \Bigg\|-K_1\Big(\mathcal{G}_0(\bar{x}_s)-\sum_{i=1}^{N}\mathcal{G}_i(\bar{x}_s)\beta_i\Big)^T\big(\nabla\phi_{c0}^T(\bar{x}_s)\tilde{\omega}_{c0}+\nabla\epsilon_{c0}(\bar{x}_s)\big)+K_2\sum_{i=1}^{N}\alpha_i\mathcal{M}_i^{-1}\mathcal{G}_i(\bar{x}_s)\cdot\big(\nabla\phi_{ci}^T(\bar{x}_s)\tilde{\omega}_{ci}+\nabla\epsilon_{c0}(\bar{x}_s)\big)\Bigg\|^2. \quad (66)$$

We can further rewrite (66) as

$$\begin{aligned}\kappa_{\vartheta_0} &\le 2\|K_1\|^2\Big\|\mathcal{G}_0(\bar{x}_s)-\sum_{i=1}^{N}\mathcal{G}_i(\bar{x}_s)\beta_i\Big\|^2\Big\|\nabla\phi_{c0}^T(\bar{x}_s)\tilde{\omega}_{c0}+\nabla\epsilon_{c0}(\bar{x}_s)\Big\|^2+2\|K_2\|^2\Big\|\sum_{i=1}^{N}\alpha_i\mathcal{M}_i^{-1}\mathcal{G}_i(\bar{x}_s)\cdot\big(\nabla\phi_{ci}^T(\bar{x}_s)\tilde{\omega}_{ci}+\nabla\epsilon_{c0}(\bar{x}_s)\big)\Big\|^2\\ &\le 4\|K_1\|^2(N+1)\mathcal{G}_M^2\Big(1+\sum_{i=1}^{N}\|\beta_i\|^2\Big)\big(\mathscr{B}_{\phi_{c0}}^2\|\tilde{\omega}_{c0}\|^2+\mathscr{B}_{\epsilon_{c0}}\big)+4\|K_2\|^2N\mathcal{G}_M^2\sum_{i=1}^{N}\|\alpha\mathcal{M}^{-1}\|^2\big(\mathscr{B}_{\phi_{ci}}^2\|\tilde{\omega}_{ci}\|^2+\mathscr{B}_{\epsilon_{ci}}\big).\end{aligned} \quad (67)$$

Considering the fact $\|\tilde{\omega}_{cj}\|^2 = \|\omega_{cj}^*-\hat{\omega}_{cj}\|^2 \le 2\|\omega_{cj}^*\|^2+2\|\hat{\omega}_{cj}\|^2 \le 4\mathscr{B}_{\omega_{cj}}^2$, we have $(\mathscr{B}_{\phi_{cj}}^2\|\tilde{\omega}_{cj}\|^2+\mathscr{B}_{\epsilon_{cj}}^2) \le (4\mathscr{B}_{\phi_{cj}}^2\mathscr{B}_{\omega_{cj}}^2+\mathscr{B}_{\epsilon_{cj}}^2)\doteq\delta_j^2, j\in\mathcal{P}$. Therefore, substituting (67) into (65), we obtain

$$\kappa \le 2\mathscr{L}_\vartheta^2\|e_s\|^2+2\Theta_M^2 \quad (68)$$

where

$$\Theta_M^2 = \Big(1+2\sum_{i=1}^{N}\|\beta_i\|^2\Big)\Big(4\|K_1\|^2(N+1)\mathcal{G}_M^2\Big(1+\sum_{i=1}^{N}\|\beta_i\|^2\Big)\delta_0^2+4\|K_2\|^2N\mathcal{G}_M^2\sum_{i=1}^{N}\|\alpha\mathcal{M}^{-1}\|^2\delta_i^2\Big)+\mathcal{G}_M^2\sum_{i=1}^{N}\left\|\mathcal{M}_i^{-1}\right\|^2\delta_i^2. \quad (69)$$

Combining (62) with (68), and setting $\mathscr{B}_{T_j}^2 = \mathscr{B}_{\phi_{cj}}^2\mathscr{B}_{\omega_{cj}}^2+\mathscr{B}_{\epsilon_{cj}}^2$, we have the first derivative of L_{1j} as

$$\dot{L}_{1j} \le -\lambda_{\min}(Q_j)||x||^2+\mathscr{B}_{T_j}^2+(N+1)\mathcal{G}_M^2\mathscr{L}_\vartheta^2\|e_s\|^2+(N+1)\mathcal{G}_M^2\Theta_M^2. \quad (70)$$

Then, for the second term $\dot{L}_{3j}$ in (59), we derive

$$\dot{L}_{3j} = -\eta_{cj}^{-1}\mathrm{tr}\Big(-\eta_{cj}\tilde{\omega}_{cj}^T\gamma_j\gamma_j^T\tilde{\omega}_{cj}+\eta_{cj}\tilde{\omega}_{cj}^T\gamma_j\frac{\Delta_{H_j}}{\psi_j}\Big). \quad (71)$$

Utilizing Young's inequality, we have

$$\dot{L}_{3j} \le -\|\gamma_j\|^2\|\tilde{\omega}_{cj}\|^2+\frac{1}{2}\eta_{cj}^{-1}\Big(\eta_{cj}^2\|\gamma_j\|^2\|\tilde{\omega}_{cj}\|^2+\Big\|\frac{\Delta_{H_j}}{\psi_j}\Big\|^2\Big) \le -\Big(1-\frac{\eta_{cj}}{2}\Big)\|\gamma_j\|^2\|\tilde{\omega}_{cj}\|^2+\frac{1}{2\eta_{cj}}\mathscr{B}_{\Delta_j}^2. \quad (72)$$

Insert (72) and (70) into (59). With some calculation, we obtain

$$\dot{L}_{cl_j} \le -a_j^2\lambda_{\min}(Q_j)||x||^2-\Big[(1-a_j^2)\lambda_{\min}(Q_j)||x||^2-(N+1)\mathcal{G}_M^2\mathscr{L}_\vartheta^2\|e_s\|^2\Big]-\Big(1-\frac{\eta_{cj}}{2}\Big)\|\gamma_j\|^2\|\tilde{\omega}_{cj}\|^2+\mathscr{B}_{T_j}^2+(N+1)\mathcal{G}_M^2\Theta_M^2+\frac{1}{2\eta_{cj}}\mathscr{B}_{\Delta_j}^2. \quad (73)$$

Hence, if the triggering condition (30) holds, we have $\dot{L}_{cl_j}\le 0$ given that $x\notin\Omega_x$ and $\tilde{\omega}_{cj}\notin\Omega_{\tilde{\omega}_{cj}}$, where

$$\Omega_x = \left\{x:\|x\|\le\sqrt{\frac{\mathscr{B}_{T_j}^2+(N+1)\mathcal{G}_M^2\Theta_M^2}{a_j^2\lambda_{\min}(Q_j)}}\right\}$$

$$\Omega_{\tilde{\omega}_{cj}} = \left\{\tilde{\omega}_{cj}:\|\tilde{\omega}_{cj}\|\le\sqrt{\frac{\mathscr{B}_{\Delta_j}^2/2\eta_{cj}}{(1-\eta_{cj}/2)\|\gamma_j\|^2}}\right\}.$$

It follows $\dot{L}_{cl}\le 0$ as long as $x\notin\Omega_x$ and $\tilde{\omega}_{cj}\notin\Omega_{\tilde{\omega}_{cj}}$. This demonstrates that x and $\tilde{\omega}_{cj}$ are guaranteed to be UUB in the continuous dynamic phase.

Now, we consider the boundedness of the jump dynamics. When $t=t_{s+1}$, the first difference of the Lyapunov function (58) is provided as follows:

$$\Delta L_{cl} = \sum_{j=0}^{N}\Big(V_j^*(x)-V_j^*(x^-)\Big)+\sum_{j=0}^{N}\Big(V_j^*(\bar{x}_{s+1})-V_j^*(\bar{x}_s)\Big)+\sum_{j=0}^{N}\eta_{cj}^{-1}\Big(\mathrm{tr}\Big(\tilde{\omega}_{cj}^T\tilde{\omega}_{cj}\Big)-\mathrm{tr}\Big(\tilde{\omega}_{cj}^{-T}\tilde{\omega}_{cj}^-\Big)\Big) \quad (74)$$

Authorized licensed use limited to: Florida Atlantic University. Downloaded on February 18,2026 at 16:15:36 UTC from IEEE Xplore. Restrictions apply.

where $x^- = x(t_{s+1}^-) = \lim_{\tau_1 \to 0^-} x(t_{s+1} + \tau_1)$, $\tilde{\omega}_{cj}^- = \tilde{\omega}_{cj}(t_{s+1}^-) = \lim_{\tau_1 \to 0^-} \tilde{\omega}_{cj}(t_{s+1} + \tau_1)$ and $\tau_1 \in (t_s - t_{s+1}, 0)$. Since we prove that $\dot{L}_{cl} < 0$ if $x \notin \Omega_x$ and $\tilde{\omega}_{cj} \notin \Omega_{\tilde{\omega}_{cj}}$ for all $t \in [t_s, t_{s+1})$, then L_{cl} is monotonically decreasing for $t \in [t_s, t_{s+1})$. Therefore, $V_j^*(x) \leq V_j^*(x^-)$ and $\mathrm{tr}(\tilde{\omega}_{cj}^T \tilde{\omega}_{cj}) \leq \mathrm{tr}(\tilde{\omega}_{cj}^{-T} \tilde{\omega}_{cj}^-)$ at the jump instants. On the other hand, for the sampled data, because we have proved that the state estimation error is UUB, then $V_j^*(\bar{x}_{s+1}) \leq V_j^*(\bar{x}_s)$. Hence, we have $\Delta L_{cl} < 0$, which indicates x and $\tilde{\omega}_{cj}$ are also UUB in the jump dynamic phase. This completes the proof. ■

IV. Experiment Studies

Example 1: Consider the following nonlinear hierarchical game with four players (open loop unstable)

$$\dot{x} = f(x) + g_0(x)u_0 + k_0(x)d_0(x) + \sum_{i=1}^{3} g_i(x)u_i + \sum_{i=1}^{3} k_i(x)d_i(x). \tag{75}$$

This game follows an asymmetric decision-making structure, where player 0 acts as the leader, making decisions first, while the followers ($i \in \{1, 2, 3\}$) respond optimally based on the leader's choice. The system dynamics are provided as

$$f(x) = \begin{bmatrix} -x_1 + x_2 + \frac{1}{2}x_1^2 x_2 \\ -x_1 - x_2 + x_1 x_2^2 + \frac{1}{4}x_2\left((\cos(2x_1) + 2)^2 + (\sin(4x_1^2) + 2)^2\right) \end{bmatrix}$$

$$g_0(x) = \begin{bmatrix} 0 \\ \cos(x_1) \end{bmatrix}, \quad g_1(x) = \begin{bmatrix} 0 \\ \cos(2x_1 + 1) \end{bmatrix}$$

$$g_2(x) = \begin{bmatrix} 0 \\ \sin(x_1 + 2) \end{bmatrix}, \quad g_3(x) = \begin{bmatrix} 0 \\ \sin(4x_1^2) + 1 \end{bmatrix}$$

where $x = [x_1, x_2]^T \in \mathbb{R}^2$ is the state, $u_0 \in \mathbb{R}$ and $u_i \in \mathbb{R}$ are the policies controlled by leader and follower i, respectively.

Furthermore, the term $k_j(x)d_j(x)$, $j \in \{0, 1, 2, 3\}$, is the unknown mismatched uncertainties applied on each player with

$$d_j(x) = \lambda_1 x_1 \cos\left(\frac{1}{x_2 + \lambda_2}\right) + \lambda_3 x_2 \sin\left(\lambda_4 x_1 x_2\right),$$

$$k_0(x) = \begin{bmatrix} \sin(x_1^2) \\ 0 \end{bmatrix}, \quad k_1(x) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$k_2(x) = \begin{bmatrix} \sin(x_1 + 1) \\ 0 \end{bmatrix}, \quad k_3(x) = \begin{bmatrix} \frac{1}{2} \\ 0 \end{bmatrix}$$

and $\lambda_1 \in [-1, 1]$, $\lambda_2 \in [-100, 0) \bigcup (0, 100]$, $\lambda_3 \in [-1, 1]$, and $\lambda_4 \in [-100, 100]$ are the unknown parameters. We can easily verify that Assumption 2 holds in this system, i.e., $\mathrm{rank}(g_j(x)) = 1 < 2$ and $g_j^T(x)k_j(x) = 0$. Therefore, the designed event-triggered RL control approach is applied for this robust control problem.

Based on Theorem 1, we first transform this hierarchical mismatched uncertain system (75) into an optimization design of Stackelberg–Nash game

$$\dot{x} = f(x) + \sum_{j=0}^{3} \mathcal{G}_j(x)\vartheta_j \tag{76}$$

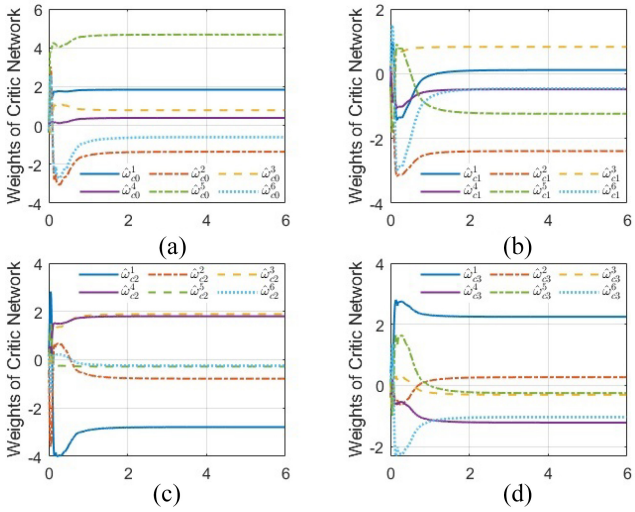

Fig. 2. Convergence process of the critic network weights $\hat{\omega}_{cj}$, $j \in \{0, 1, 2, 3\}$: (a) Leader, (b) Follower 1, (c) Follower 2, and (c) Follower 3.

where $\vartheta_j = [u_j, v_j]^T$, $v_j \in \mathbb{R}$ is an auxiliary input and $\mathcal{G}_j(x) = [g_j(x), k_j(x)]$, $j \in \{0, 1, 2, 3\}$. Note that the open-loop configuration of (76) is also unstable.

Since $||d_j(x)|| \leq ||x||$, define the upper bounds as $\mathscr{D}_d(x) = 4||x||$ and $\mathscr{B}_v^2(x) = ||\sum_{j=0}^{3} \rho_j k_j(x) v_j||^2$ with $\rho_j = 2$. Design the utility function based on (6) for leader and based on (7) for followers with $Q_j = 5I_2$, $\mathcal{M}_j = I_2$, $j \in \{0, 1, 2, 3\}$, and $\alpha_i = 0.2I_2$, $\beta_i = 0.2I_2$, $i \in \{1, 2, 3\}$.

Then, establish the critic network for each player to approximate the performance index $\hat{V}_j(x)$ based on (49) and calculate the event-triggered control policy $\vartheta_j(\bar{x}_s) = [u_j(\bar{x}_s), v_j(\bar{x}_s)]^T$ based on (51) and (52). The learning rate is chosen as $\eta_{cj} = 0.05$. The critic network for each player is designed as a three layer network with the input as $C_j = [x_1, x_2, u_j, v_j]^T$ and the output as $\hat{V}_j(x)$. We consider 6 neurons for the hidden layer with the activation function as $\phi_j(x) = ([1 - e^{-\omega_{c1,j}^T C_j}]/[1 + e^{-\omega_{c1,j}^T C_j}])$, where $\omega_{c1,j}$ are the weights between the input and hidden layer. In this article, we randomly choose $\omega_{c1,j} \in [-0.5, 0.5]$ at the beginning and fix the values thereafter. Then, the weights between the output and hidden layer $\hat{\omega}_{cj}$ are adjusted based on (55).

Select the sample interval as $0.01s$ and let the initial state be $x(0) = [1, -0.5]^T$. Choose $a_j = 0.5$ and $\mathcal{L}_{\vartheta_j} = 3$, $j \in \{0, 1, 2, 3\}$, for the triggering threshold (30). Besides, to satisfy the persistent excitation condition, we add a probing noise $0.1\sin^2(t)\cos(t) + 0.1\sin^2(2t)\cos(0.1t)$ to the control $u_j(\bar{x}_s)$ for the first 80 time steps.

The learning process spans a total of 6 s (600 time steps × 0.01 s sample interval). The learning evolution of the critic network weights between the hidden and output layer $\hat{\omega}_{cj}$ is provided in Fig. 2. We can observe that all the weights can converge quickly, which demonstrates the optimal learning process of the developed method. Fig. 3 provides the event-triggered control policy $u_j(\bar{x}_s)$ and event-triggered auxiliary policy $v_j(\bar{x}_s)$, $j \in \{0, 1, 2, 3\}$. The updating steps can be detected, which confirms that both control signals are intermittent feedback and updated aperiodically. Based on the

Authorized licensed use limited to: Florida Atlantic University. Downloaded on February 18,2026 at 16:15:36 UTC from IEEE Xplore. Restrictions apply.

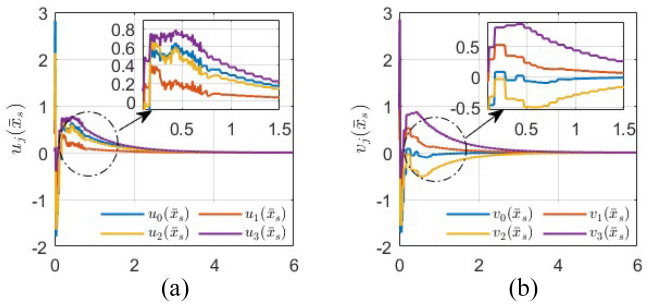

Fig. 3. Evolution of event-triggered control policies $u_j(\bar{x}_s)$ and $v_j(\bar{x}_s)$ in the learning process. (a) t/s. (b) t/s.

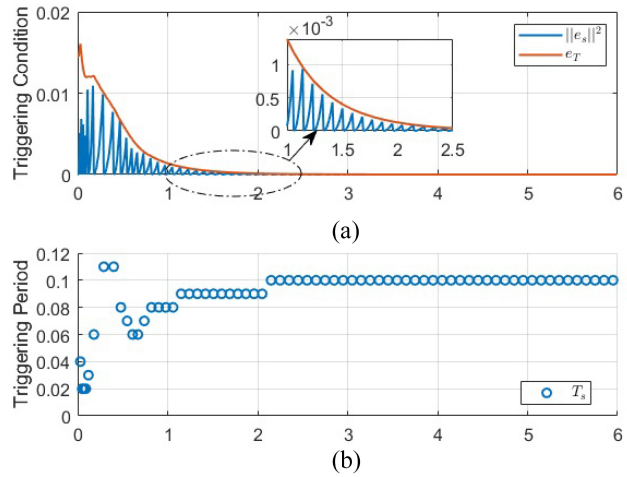

Fig. 4. (a) Triggering condition verification. (b) Intersampling period.

definition, $\vartheta_j(\bar{x}_s)$ is constituted by $\vartheta_j(\bar{x}_s) = [u_j(\bar{x}_s), v_j(\bar{x}_s)]^T$. The triggering condition and triggering period are shown in Fig. 4. Particularly, Fig. 4(a) compares the square norm of e_s with the triggering threshold e_T, ensuring that $||e_s||^2$ is strictly smaller than e_T which satisfies the triggering condition (30) and guarantees the asymptotic stability of the system. In addition, Fig. 4(b) provides the triggering period $T_s = t_{s+1} - t_s$, $s \in \mathbb{N}$. We can observe that $T_{s,\min} = 0.02$ which verifies that the Zeno behavior does not occur. Note that T_s becomes constant after 2.2 seconds. This indicates that the learning process has reached the optimal solution. Furthermore, we compare the triggering number in the developed event-triggered method with the time-triggered control method in Fig. 5. It is shown that the developed method only needs 66 state samples while the traditional time-triggered method requires 600 state samples, which is 89% reduction of the control updates. This indicates that the developed method can effectively reduce the computational burden.

After the training, we evaluate the control performance of the obtained event-triggered control policy on the mismatched uncertain hierarchical system (75). Without loss of generality, we randomly select admissible unknown parameters of uncertainty and measure the mean square error (MSE) of the state $x = [x_1, x_2]^T$ for each set of parameters. We perform 5000 independent runs and the results are presented in Fig. 6. It is demonstrated that the MSE of the state remains finite for all uncertain parameters under the designed event-triggered

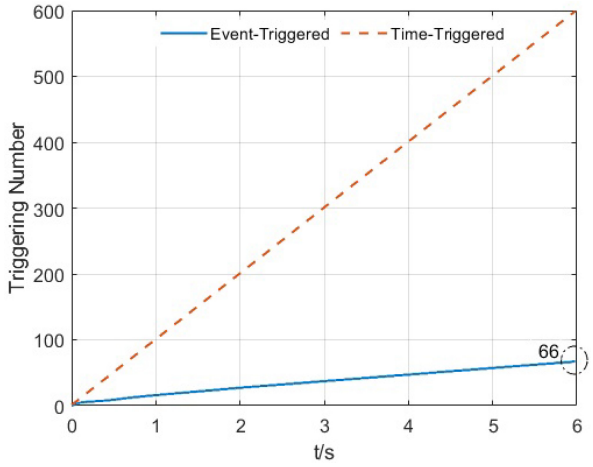

Fig. 5. Comparison of triggering number.

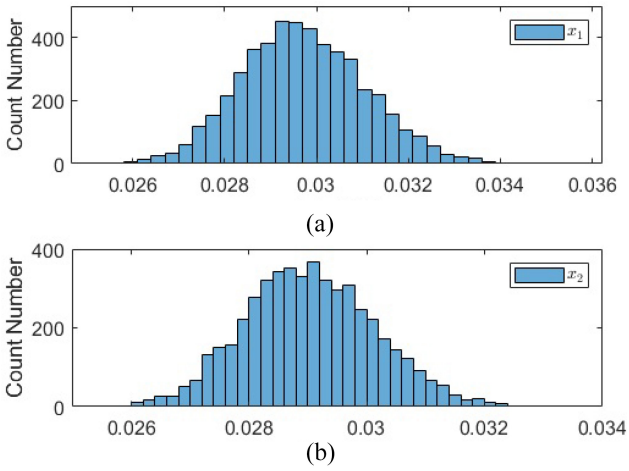

Fig. 6. Histogram of MSE for state in robust control process: (a) MSE for x_1, and (b) MSE for x_2.

control policy. This implies that the closed-loop system is compelled to be asymptotically stable and exhibits robustness against any admissible uncertainties. Fig. 7 shows a general state trajectory in the robust control process. These results further demonstrate the stability preservation of the designed robust-optimal transformation, ensuring that the transformed system accurately reflects the original dynamics. Moreover, the results validate the effectiveness of our developed method in achieving reliable and robust performance in hierarchical multiplayer systems, even under complex uncertainties.

Example 2: To further evaluate the proposed method, we apply it to a ten-player hierarchical system with the following:

$$\dot{x} = f(x) + g_0(x)u_0 + k_0(x)d_0(x) + \sum_{i=1}^{9} g_i(x)u_i + \sum_{i=1}^{9} k_i(x)d_i(x) \tag{77}$$

where $x = [x_1, x_2]^T \in \mathbb{R}^2$ is the state, $u_0 \in \mathbb{R}$ and $u_i \in \mathbb{R}$ are the policies controlled by leader and follower $i \in \{1, 2, 3, \ldots, 9\}$, respectively. The dynamics are given as

$$f(x) = \begin{bmatrix} -x_1 + x_2 + \sin(x_1^2) \\ -x_1 - x_2 + x_1 x_2^2 + 0.5x_2(\cos(2x_1) + 2)^2 \end{bmatrix} \tag{78}$$

Authorized licensed use limited to: Florida Atlantic University. Downloaded on February 18,2026 at 16:15:36 UTC from IEEE Xplore. Restrictions apply.

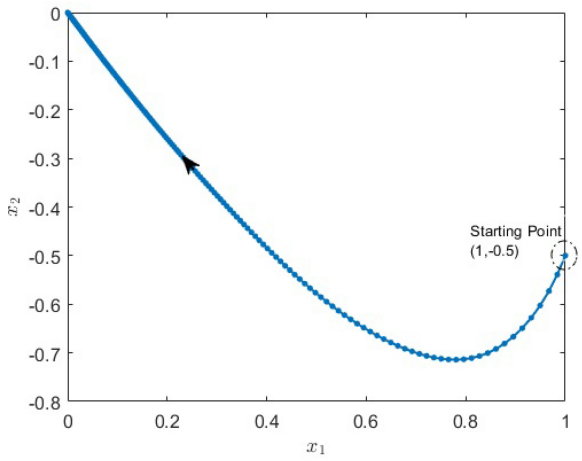

Fig. 7. General state trajectory in robust control process.

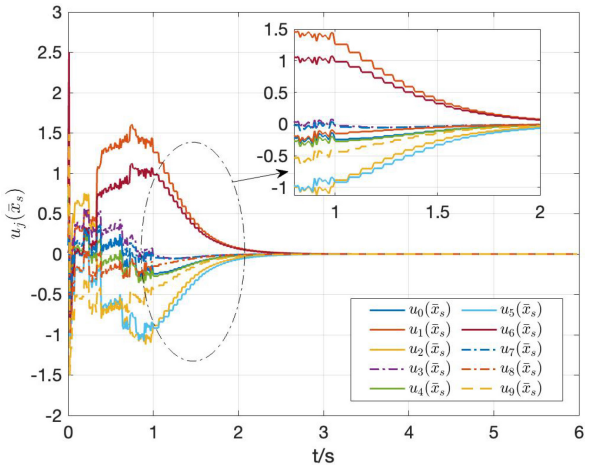

Fig. 8. Evolution of event-triggered control policy $u_j(\bar{x}_s), j \in \{0, 2, 3, \ldots, 9\}$.

and $g_0(x) = [0, \cos(x_1)]^T$, $g_1(x) = [0, 1]^T$, $g_2(x) = [0, \cos(2x_1)]^T$, $g_3(x) = [0, 2]^T$, $g_4(x) = [0, \sin(4x_1) + 1]^T$, $g_5(x) = [2\cos(x_1), 0]^T$, $g_6(x) = [2\sin(x_2), 0]^T$, $g_7(x) = [0, \cos(x_1)]^T$, $g_8(x) = [0, \sin(x_2)]^T$, and $g_9(x) = [0, \sin(2x_1) + 1]^T$. Here, the leader 0 acts first and the follower $i \in \{1, 2, 3, \ldots, 9\}$ respond to the leader's decision.

The unknown mismatched uncertainties $k_j(x)d_j(x)$ are applied on each player $j \in \{0, 1, 2, \ldots, 9\}$ with $k_0(x) = [\sin(x_1^2), 0]^T$, $k_1(x) = [\cos(2x_2^2), 0]^T$, $k_2(x) = [\sin(x_2 + 2), 0]^T$, $k_3(x) = [1, 0]^T$, $k_4(x) = [\sin(x_1 + 1), 0]^T$, $k_5(x) = [0, \sin(2x_2 + 1)]^T$, $k_6(x) = [0, \cos x_1^2]^T$, $k_7(x) = [1/2, 0]^T$, $k_8(x) = [1, 0]^T$, $k_9(x) = [\sin(x_1), 0]^T$, and $d_j(x) = p_1 x_1 \cos(x_2) + p_2 x_2 \sin(x_1)$, where $p_1, p_2 \in [-2, 2]$ are unknown parameters.

We implement the designed event-triggered RL control method to address this hierarchical robust control problem. The auxiliary plant is constructed based on (4) with $N = 9$. The utility function is designed in (6) for the leader and in (7) for the followers with $Q_j = 8I_2$, $\mathcal{M}_j = 2I_2$, $j = \{0, 1, 2, \ldots, 9\}$, and the coupling coefficients $\alpha_i = 0.1I_2$, $\beta_i = 0.1I_2$, $i \in \{1, 2, 3, \ldots, 9\}$. The triggering threshold is established with $a_j = 0.5$ and $\mathcal{L}_{\vartheta_j} = 3$. The learning rate process spans a total of 6 s (600 time steps $\times$ 0.01 s sample interval). To satisfy the persistent excitation condition, the probing noise is added to the control $u_j(\bar{x}_s)$ for the first 100 time steps.

The event-triggered control policy $u_j(\bar{x}_s)$ and the event-triggered auxiliary policy $v_j(\bar{x}_s)$ are provided in Fig. 8 and Fig. 9, respectively. We can observe that both control signals are intermittent feedback and updated aperiodically, adjusting updates based on system conditions rather than fixed time intervals. The triggering condition and triggering period are illustrated in Fig. 10. As shown in Fig. 10(a), the term $||e_s||^2$ consistently remains below the threshold e_T to ensure the stability of the learning process. This indicates that the learning updates are effectively regulated, preventing excessive or unnecessary updates while maintaining control performance. Furthermore, the triggering period in Fig. 10(b) confirms that Zeno behavior is successfully avoided. Note that this method triggers only 134 control updates comparing with 600 updates in traditional time-triggered methods.

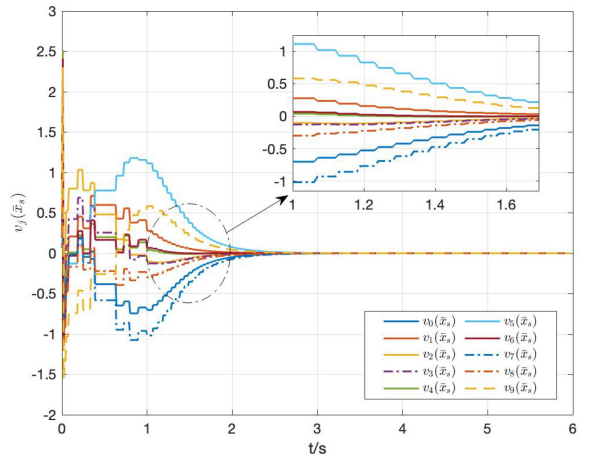

Fig. 9. Evolution of event-triggered control policy $v_j(\bar{x}_s), j \in \{0, 2, 3, \ldots, 9\}$.

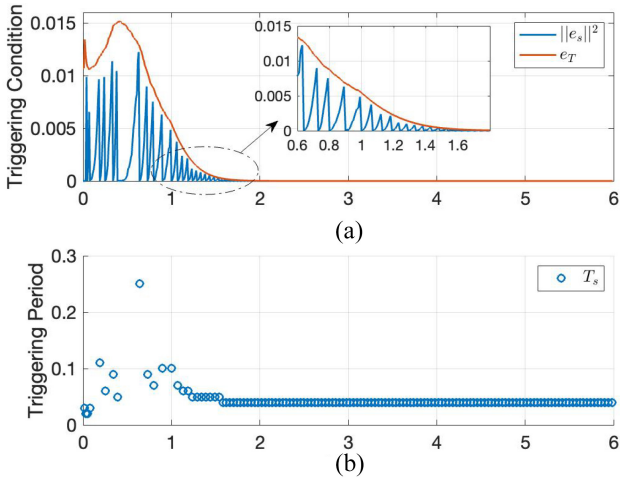

Fig. 10. (a) Triggering condition verification. (b) Intersampling period.

After that, we fix the critic network weights and apply the designed feedback controller to the original uncertain system (77). Set the uncertain parameters to $p_1 = -2$ and $p_2 = 2$. We conduct the robust control process for 500 time steps $\times$ 0.01 sample interval $= 5s$, starting from the

Authorized licensed use limited to: Florida Atlantic University. Downloaded on February 18,2026 at 16:15:36 UTC from IEEE Xplore. Restrictions apply.

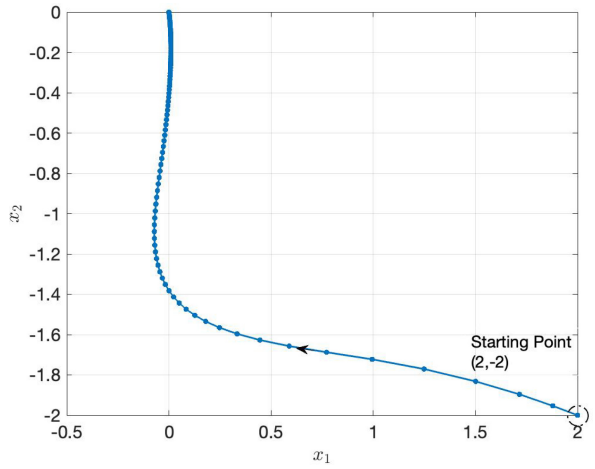

Fig. 11. State trajectory in robust control process.

initial state $[2, -2]^T$. The corresponding state trajectory under this control scheme is shown in Fig. 11. We can observe that the system rapidly converges to the equilibrium point under the designed controller. These results further validate our developed learning-based control method and demonstrate its capability to handle hierarchical multiplayer systems with mismatched uncertainties while ensuring stability and performance.

V. Conclusion

This article develops an RL-based robust event-triggered approach for hierarchical multiplayer systems with mismatched uncertainties. The proposed framework addresses the computational complexities and mismatched uncertainties inherent in the learning-based asymmetric decision-making process. Specifically, the problem is transformed into the Stackelberg–Nash game framework, complemented by an intelligent event-triggered hierarchical robust-optimal approach developed with RL techniques. Theoretical analysis is provided to demonstrate that the solution of the transformed event-triggered control scheme ensures the asymptotic stabilization of the original hierarchical robust control problem. Neural network techniques are applied to implement the developed method and the detailed stability guarantee is discussed. Finally, the numerical studies are conducted to validate the effectiveness of the proposed method.

References

[1] D. Xie and X. Zhong, "Semicentralized deep deterministic policy gradient in cooperative starcraft games," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 33, no. 4, pp. 1584–1593, Apr. 2022.

[2] J. Perolat et al., "Mastering the game of stratego with model-free multiagent reinforcement learning," *Science*, vol. 378, no. 6623, pp. 990–996, 2022.

[3] H. Huang, Z. Hu, Z. Lu, and X. Wen, "Network-scale traffic signal control via multiagent reinforcement learning with deep spatiotemporal attentive network," *IEEE Trans. Cybern.*, vol. 53, no. 1, pp. 262–274, Jan. 2023.

[4] K. Shao, Y. Zhu, and D. Zhao, "Starcraft micromanagement with reinforcement learning and curriculum transfer learning," *IEEE Trans. Emerg. Topics Comput. Intell.*, vol. 3, no. 1, pp. 73–84, Feb. 2019.

[5] G. Cui, Q.-S. Jia, and X. Guan, "Energy management of networked microgrids with real-time pricing by reinforcement learning," *IEEE Trans. Smart Grid*, vol. 15, no. 1, pp. 570–580, Jan. 2024.

[6] Y. Li et al., "Multiagent reinforcement learning-based signal planning for resisting congestion attack in green transportation," *IEEE Trans. Green Commun. Netw.*, vol. 6, no. 3, pp. 1448–1458, Sep. 2022.

[7] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*. Cambridge, MA, USA: MIT Press, 2018.

[8] D. Silver et al., "Mastering the game of go with deep neural networks and tree search," *Nature*, vol. 529, no. 7587, pp. 484–489, 2016.

[9] D. Silver et al., "Mastering the game of go without human knowledge," *Nature*, vol. 550, no. 7676, pp. 354–359, 2017.

[10] 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.

[11] D. Liu, H. Liu, J. Lü, and F. L. Lewis, "Time-varying formation of heterogeneous multiagent systems via reinforcement learning subject to switching topologies," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 70, no. 6, pp. 2550–2560, Jun. 2023.

[12] K. G. Vamvoudakis, H. Modares, B. Kiumarsi, and F. L. Lewis, "Game theory-based control system algorithms with real-time reinforcement learning: How to solve multiplayer games online," *IEEE Control Syst. Mag.*, vol. 37, no. 1, pp. 33–52, Feb. 2017.

[13] X. Zhong and H. He, "GrHDP solution for optimal consensus control of multiagent discrete-time systems," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 50, no. 7, pp. 2362–2374, Jul. 2020.

[14] Y. Yang, H. Modares, K. G. Vamvoudakis, and F. L. Lewis, "Cooperative finitely excited learning for dynamical games," *IEEE Trans. Cybern.*, vol. 54, no. 2, pp. 797–810, Feb. 2024.

[15] H. Zhang, Y. Cai, Y. Wang, and H. Su, "Adaptive bipartite event-triggered output consensus of heterogeneous linear multiagent systems under fixed and switching topologies," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 31, no. 11, pp. 4816–4830, Nov. 2020.

[16] T. T. Nguyen, N. D. Nguyen, and S. Nahavandi, "Deep reinforcement learning for multiagent systems: A review of challenges, solutions, and applications," *IEEE Trans. Cybern.*, vol. 50, no. 9, pp. 3826–3839, Sep. 2020.

[17] J. Zhang, H. Zhang, and S. Sun, "Adaptive dynamic event-triggered bipartite time-varying output formation tracking problem of heterogeneous multiagent systems," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 54, no. 1, pp. 12–22, Jan. 2024.

[18] J. Chai, W. Li, Y. Zhu, D. Zhao, Z. Ma, K. Sun, and J. Ding, "UNMAS: Multiagent reinforcement learning for unshaped cooperative scenarios," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 34, no. 4, pp. 2093–2104, Apr. 2023.

[19] Q. Wei, X. Wang, X. Zhong, and N. Wu, "Consensus control of leader-following multi-agent systems in directed topology with heterogeneous disturbances," *IEEE/CAA J. Automatica Sinica*, vol. 8, no. 2, pp. 423–431, Feb. 2021.

[20] Q. Wei, Y. Li, J. Zhang, and F.-Y. Wang, "VGN: Value decomposition with graph attention networks for multiagent reinforcement learning," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 35, no. 1, pp. 182–195, Jan. 2024.

[21] P. Paudyal, P. Munankarmi, Z. Ni, and T. M. Hansen, "A hierarchical control framework with a novel bidding scheme for residential community energy optimization," *IEEE Trans. Smart Grid*, vol. 11, no. 1, pp. 710–719, Jan. 2020.

[22] M. Yu, J. Jiang, X. Ye, X. Zhang, C. Lee, and S. H. Hong, "Demand response flexibility potential trading in smart grids: A multileader multifollower Stackelberg game approach," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 53, no. 5, pp. 2664–2675, May 2023.

[23] Y. Yang, W. Wang, L. Liu, K. Dev, and N. M. F. Qureshi, "AoI optimization in the UAV-aided traffic monitoring network under attack: A Stackelberg game viewpoint," *IEEE Trans. Intell. Transp. Syst.*, vol. 24, no. 1, pp. 932–941, Jan. 2023.

[24] N. Groot, B. De Schutter, and H. Hellendoorn, "Toward system-optimal routing in traffic networks: A reverse Stackelberg game approach," *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 1, pp. 29–40, Feb. 2015.

[25] P. Hang, C. Lv, Y. Xing, C. Huang, and Z. Hu, "Human-like decision making for autonomous driving: A noncooperative game theoretic approach," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 4, pp. 2076–2087, Apr. 2021.

[26] T. Başar and G. J. Olsder, *Dynamic Noncooperative Game Theory*. Philadelphia, PA, USA: SIAM, 1998.

Authorized licensed use limited to: Florida Atlantic University. Downloaded on February 18,2026 at 16:15:36 UTC from IEEE Xplore. Restrictions apply.

[27] K. G. Vamvoudakis, F. L. Lewis, and W. E. Dixon, "Open-loop Stackelberg learning solution for hierarchical control problems," *Int. J. Adapt. Control Signal Process.*, vol. 33, no. 2, pp. 285–299, 2019.
[28] Y. Huang and J. Zhao, "Active interdiction defence scheme against false data-injection attacks: A Stackelberg game perspective," *IEEE Trans. Cybern.*, vol. 54, no. 1, pp. 162–172, Jan. 2024.
[29] R. Yu, Y.-H. Chen, and Q. Wang, "A Stackelberg game-theoretic exploration rendering robustness and optimality for performance improvement of fuzzy mechanical systems," *IEEE Trans. Cybern.*, vol. 53, no. 1, pp. 289–302, Jan. 2023.
[30] C. Mu, K. Wang, Q. Zhang, and D. Zhao, "Hierarchical optimal control for input-affine nonlinear systems through the formulation of Stackelberg game," *Inf. Sci.*, vol. 517, pp. 1–17, May 2020.
[31] H. Zhong, Z. Yang, Z. Wang, and M. I. Jordan, "Can reinforcement learning find Stackelberg-Nash equilibria in general-sum Markov games with myopically rational followers?" *J. Mach. Learn. Res.*, vol. 24, no. 35, pp. 1–52, 2023.
[32] J. Moon and T. Başar, "Linear quadratic mean field Stackelberg differential games," *Automatica*, vol. 97, pp. 200–213, Nov. 2018.
[33] H. Mukaidani and H. Xu, "Stackelberg strategies for stochastic systems with multiple followers," *Automatica*, vol. 53, pp. 53–59, Mar. 2015.
[34] M. Li, J. Qin, Q. Ma, W. X. Zheng, and Y. Kang, "Hierarchical optimal synchronization for linear systems via reinforcement learning: A Stackelberg–Nash game perspective," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 4, pp. 1600–1611, Apr. 2021.
[35] M. Li, J. Qin, N. M. Freris, and D. W. Ho, "Multiplayer Stackelberg–Nash game for nonlinear system via value iteration-based integral reinforcement learning," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 33, no. 4, pp. 1429–1440, Apr. 2022.
[36] M. R. Satouri, H. Kebriaei, A. Razminia, and M. J. Yazdanpanah, "Robust on-line ADP-based solution of a class of hierarchical nonlinear differential game," 2019, *arXiv:1907.11414*.
[37] X. Zhong and H. He, "An event-triggered ADP control approach for continuous-time system with unknown internal states," *IEEE Trans. Cybern.*, vol. 47, no. 3, pp. 683–694, Mar. 2017.
[38] F. Zhao, S. Luo, W. Gao, and C. Wen, "Event-triggered cooperative adaptive optimal output regulation for multiagent systems under switching network: An adaptive dynamic programming approach," *IEEE Trans. Syst., Man, Cybern. Syst.*, vol. 55, no. 3, pp. 1707–1721, Mar. 2025.
[39] X. Yang and H. He, "Adaptive critic learning and experience replay for decentralized event-triggered control of nonlinear interconnected systems," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 50, no. 11, pp. 4043–4055, Nov. 2020.
[40] T. Li, D. Yang, X. Xie, and H. Zhang, "Event-triggered control of nonlinear discrete-time system with unknown dynamics based on HDP (λ)," *IEEE Trans. Cybern.*, vol. 52, no. 7, pp. 6046–6058, Jul. 2022.
[41] K. G. Vamvoudakis, "Event-triggered optimal adaptive control algorithm for continuous-time nonlinear systems," *IEEE/CAA J. Automatica Sinica*, vol. 1, no. 3, pp. 282–293, Jul. 2014.
[42] C. Qin, T. Zhu, K. Jiang, and Y. Wu, "Integral reinforcement learning-based dynamic event-triggered safety control for multiplayer Stackelberg–Nash games with time-varying state constraints," *Eng. Appl. Artif. Intell.*, vol. 133, Jul. 2024, Art. no. 108317.
[43] W. Song, J. Feng, H. Zhang, and W. Wang, "Dynamic event-triggered formation control for heterogeneous multiagent systems with nonautonomous leader agent," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 34, no. 12, pp. 9685–9699, Dec. 2022.
[44] X. Zhong and H. He, "Event-triggered multi-agent optimal regulation using adaptive dynamic programming," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, 2020, pp. 1–8.
[45] M. Lin, B. Zhao, and D. Liu, "Event-triggered robust adaptive dynamic programming for multiplayer Stackelberg–Nash games of uncertain nonlinear systems," *IEEE Trans. Cybern.*, vol. 54, no. 1, pp. 273–286, Jan. 2024.

Xiangnan Zhong (Member, IEEE) is currently an Associate Professor with the Department of Electrical Engineering and Computer Science, Florida Atlantic University, Boca Raton, FL, USA. Her research interests include computational intelligence, reinforcement learning, cyber–physical systems, networked control systems, neural networks, and optimal control.

Prof. Zhong received the National Science Foundation (NSF) Faculty Early Career Development (CAREER) Award in 2021 and the NSF CRII Award in 2019. She was a recipient of the International Neural Network Society (INNS) Aharon Katzir Young Investigator Award in 2021 and the INNS Doctoral Dissertation Award in 2019. She has been serving as an Associate Editor of IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS since 2021 and IEEE INTERNET OF THINGS JOURNAL since 2023.

Zhen Ni (Senior Member, IEEE) is currently an Associate Professor with the Department of Electrical Engineering and Computer Science, Florida Atlantic University, Boca Raton, FL, USA. His research interests mainly include artificial intelligence, and computational methods, and reinforcement learning.

Dr. Ni received the Senior Faculty Teaching Award from FAU College of Engineering and Computer Science in 2024 and the NSF CAREER Award in 2021. He has been an Associate Editor of IEEE TRANSACTIONS ON EMERGING TOPICS IN COMPUTATIONAL INTELLIGENCE since 2025 and IEEE INTERNET OF THINGS JOURNAL since 2021. He served as an Associate Editor for IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS between 2019–2024, and *IEEE Computational Intelligence Magazine* between 2018–2023.

Authorized licensed use limited to: Florida Atlantic University. Downloaded on February 18,2026 at 16:15:36 UTC from IEEE Xplore. Restrictions apply.