IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 20 September 2022, accepted 8 October 2022, date of publication 14 October 2022, date of current version 21 October 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3214849

RESEARCH ARTICLE

Dynamic Event-Triggered Consensus Control for Multi-Agent Systems Using Adaptive Dynamic Programming

QI ZHANG[1], YANG YANG[1], (Member, IEEE), XIAORAN XIE[1], CHUNMING XU[2], AND HAN YANG[3]

[1]School of Electronic Information Engineering, Changchun University of Science and Technology, Changchun 130013, China
[2]Changchun Shikai Science and Technology Industry Company Ltd., Changchnu 130015, China
[3]Beijing Engineering Research Center of Emergency Survival Security, Beijing 102206, China

Corresponding author: Yang Yang (yangyang@cust.edu.cn)

This work was supported by Science and Technology Development Program of Jilin Province under Grant no.20200401066GX.

ABSTRACT The consensus optimal control problem for a class of linear multi-agent systems with directed communication networks is studied in this paper using adaptive dynamic programming. To overcome the restrictions of the agent's low processing capability and to extend the actuator's lifetime, consider the event-triggered. In the beginning, a dynamic event-triggered is provided, with several existing static event-triggered serving as special examples. When using the dynamic event-triggered, a longer interval can be shown between any two consecutive event-triggered. Designing a dynamic event-triggered control law becomes more challenging when implemented in directed networks. In addition, on the basis of dynamic event-triggered, a novel adaptive dynamic programming is used to construct a suitable dynamic event-triggered control law, which employs just the interaction information between agents and does not need model dynamics, overcoming the difficulty of solving the algebraic Riccati equation. Finally, the effectiveness of the proposed method is verified by simulation results, and no agent exhibits Zeno behavior.

INDEX TERMS Multi-agent systems, event-triggered, adaptive dynamic programming, reinforcement learning.

I. INTRODUCTION

Multi-agent systems have found significant use in physics, social sciences, biology, and engineering in recent years [1], [2], [3]. It can be used to address issues that are too challenging for a single agent to tackle. This has consequently led to an increased interest in distributed control of multi-agent systems. The consensus problem serves as the theoretical and practical foundation for agents cooperation in the multi-agent distributed control issue [4], [5]. However, many studies make the assumption that each agent designs the proper feedback control laws using the continuous signals of its neighbors. In actuality, such an assumption means that a perfect communication network would require limitless bandwidth, which is not the case [6], [7]. By allowing each agent to utilize a sampled signal rather of a continuous signal from its neighbors, sampled signals give an option to control design [8]. Typically, the agent's signal is sampled at a fixed interval of time. In cases where the time period of a continuously sampled signal are small, a high number of unnecessary sampled signals are sent between agents through the communication channel. If some devices are powered by batteries, a lot of network bandwidth and energy is squandered [9], [10].

Because it lowers the sample frequency, event-triggered sampling, an acyclic sampling technique [11], performs better than periodic sampling. The use of event-triggered sampling versus periodic sampling for an isolated system was compared in detail by [12]. Based on [13], several theoretical conclusions were drawn about input-to-state stability under event-triggered sampling. Recently, some event-triggered

The associate editor coordinating the review of this manuscript and approving it for publication was Qiang Li.

techniques were included into consensus algorithms for multi-agent systems with varying dynamics to reduce communication burden among agents. For example, multi-agent systems with first-order dynamics [15], [16], second-order dynamics [17], [18] and high-order linear dynamics [19], [20], with communication periods connected to specific event-triggered methods, were investigated [14]. More precisely, in [19], it was believed that all agents' communication networks were undirected, but in [20], [21], and [22], they were considered to be directed. Zeno behavior refers to an infinite accumulation of executions at one moment, which is an aberrant phenomena that typically manifests in a system. The triggering condition is modified so that all triggering instants must occur more than zero time intervals apart to rule out Zeno behavior. Consensus has been reached with an error in the inserted constant, which is not desirable. It was suggested in [22] that a triggering condition dependent on state exist, but the closed-loop system's lack of exhibiting Zeno behavior has just been shown. Using dynamic triggering, Girard [23] updated the results of [13] by incorporating them into linear stability analysis. Protocols for event-triggered consensus for complex dynamical networks with discrete time delay [25] and linear multiagent systems [24] introduced the concept of dynamic triggering. [26] introduces a dynamic event-triggered that is a special case of some existing static event-triggered. Using the innovative dynamic triggering, a protocol for event-triggered control is developed. Consensus can be obtained with an exponential convergence rate using this control system. However, there is a difficult challenge that dynamic triggering methods must solve: how to create a dynamic event-driven control law. The creation of adequate event-triggered control laws in prior work is dependent on solutions to several matrix inequalities, the existence of which is not easily ensured.

Current study, on the other hand, necessitates a full grasp of the dynamics of the agents, which is problematic in many actual settings. Previous research on the design of adaptive controllers for uncertain linear systems has gone into great detail. The typical method for developing adaptive optimal control laws is to first compute the algebraic Riccati equation using the system parameters. This problem necessitates exact dynamics, which is problematic since most systems in practice are too intricate, resulting in erroneous dynamics. As a result, finding a proven solution to this problem is essential.

ADP is regarded as one of the fundamental ways for achieving optimal control laws for a variety of optimal control issues because it has strong self-learning and self-adaptive capabilities and has evolved into an essential optimal control method that is similar to the brain [27]. ADP was known by several different names, including "adaptive critic designs" [28] "approximate dynamic programming" [29] and "neural dynamic programming" [30]. It includes both value and policy iteration. The value iterative ADP approach was shown to be convergence [31]. For optimal control of the continuous-time system, policy iteration is offered [32]. Continuous-time complex-valued systems have been successfully addressed via policy iteration [33]. Furthermore, convergence and stability proof were developed for discrete-time policy iteration [34]. In summary, adaptive dynamic programming has a complete theoretical foundation and is well suited to solve multi-agent systems' consensus optimal control problems. Therefore, ADP approaches have been applied on multi-agent systems to deal with the optimal distributed control problems [38], [39], [40]. Reference [38] investigates the robust optimal consensus for nonlinear multi-agent systems through the local adaptive dynamic programming (ADP) approach and the event- triggered control method. Reference [39] is concerned with the design of distributed optimal coordination control for nonlinear multi-agent systems based on event-triggered adaptive dynamic programming method. In [40], non-quadratic cost functions are introduced to handle input constraints and a novel distributed optimal consensus protocol is derived based on event-triggered adaptive dynamic programming method.

In previous work, some works have applied dynamic event-triggered to the problem of multi-agent systems' consensus, such as [41], in contrast to it, where the update of the dynamic variable relies on the dynamic variable at the last sampling moment, our dynamic variable is updated by the difference between the combined measurement variable and the measurement error. But under a dynamic triggering mechanism, there is a challenging issue to be addressed: how to design a dynamic event-triggered control law. To overcome this challenge, we propose an online ADP in this study to build a dynamic event-triggered control law for solving the optimal consensus issue of multi-agent systems, and successfully solved the ARE, avoiding the issue that matrix inequalities may not be easily guaranteed. But unlike [38], [39], [40], [41], and [42], we have the advantage of using system data instead of precise dynamics to solve consensus problems in multi-agent systems. The ADP developed in this chapter only requires an arbitrary stabilizing control policy for the entire learning phase and can effectively use the same amount of online measurements for multiple iteration steps, at the cost of a higher computational burden at a single iteration time point. So this work is of great significance.

As for the remainder of the paper, it should be organized as follows: In Section 2, there is some algebraic graph theory knowledge presented and problem formulation is derived. Section 3 shows a dynamic event-triggered, and demonstrates that it does not exhibit Zeno behavior. A large interevent time can be achieved by using the dynamic event-triggered. Unnecessarily requiring no prior knowledge of the dynamics of the system, Section 4 proposes an online ADP technique for designing a dynamic event-triggered control law. Section 5 provides an example to show our approach's effectiveness. The final section of the paper provides a brief conclusion.

II. PROBLEM FORMULATION

A. PRELIMINARIES

In this paper, graph theory is used as a powerful mathematical tool to analyze multi-agent systems. Whether the information flow is unidirectional or bidirectional, a weighted graph may explain the structure of a communication network. Assuming that M is a symmetrical real matrix of suitable size, we define $\lambda_{\min}(M) = \min_i \lambda_i(M)$ and $\lambda_{\max}(M) = \max_i \lambda_i(M)$, where $\lambda_i(M)$ is an arbitrary eigenvalue of M. Define $\operatorname{col}(x_1, \ldots, x_n) = \left[x_1^T, \ldots, x_N^T\right]^T$, where $x_i \in R^n (i = 1, \ldots, N)$. If $A \in R^{m\times n}$ and $B \in R^{p\times q}$, then $A \otimes B \in R^{mp\times nq}$, where $\otimes$ is the Kronecker product. All agents' interaction topology is directed, as shown by $\mathcal{G} = (\mathcal{V}, \mathcal{E})$. There are N agents in total. Then, and the edge set $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ is made up of all the channels that connect two agents. Particularly, when a link starts at agent-j and ends at agent-i, as shown by $(i, j) \in \mathcal{E}$. It is here that agent-j can be referred to as an agent-i's in-neighbor, and the entire set of in-neighbors can be considered in-neighbor set of agent-i, which is denoted by $\mathcal{N}_i = \{j \in \mathcal{V} \mid (i, j) \in \mathcal{E}\}$. $\mathcal{M}_i = \{j \in \mathcal{V} \mid (j, i) \in \mathcal{E}\}$ denotes the agent-i's out-neighbor set. Define $\mathcal{A}(\mathcal{G}) = (a_{ij})_{N\times N}$. In the case where j is agent i's an in-neighbor, $a_{ij} = 1$, and in the absence of this, $a_{ij} = 0$, $\mathcal{A}(\mathcal{G})$ refers to the adjacency matrix. We let $L = (l_{ij})_{N\times N}$, where $l_{ii} = \sum_{j=1, j\neq i}^{N} a_{ij}$ and $l_{ij} = -a_{ij}$, if $i \neq j$. The Laplacian matrix of $\mathcal{G}$ is denoted by L. It implies that $a_{ij} = a_{ji}, \forall i, j = 1, \ldots, N$, for an undirected graph, while this is not always the case for a directed graph. In the directed graph, there are certain ordered edges $(i_1, i_2), \ldots, (i_{k-1}, i_k)$ that make up a directed path from i_k to i_j, and in this situation, it is said that agent-i_k may reach agent-i_1. It is possible for any agent in a directed graph $\mathcal{G}$ to contact any other agent in the graph, the graph $\mathcal{G}$ is said to be strongly linked.

B. PROBLEM FORMULATION

It is shown that a class of multi-agent systems has the following dynamic of agent-i:

$$\dot{x}_i = Ax_i + Bu_i, \qquad i = 1, \ldots, N. \tag{1}$$

The n-dimensional state is represented by x_i, while the m-dimensional control input is represented by u_i, there are two system matrices with the appropriate dimensions, A and B. Consider a directed graph $\mathcal{G}$ in which all agents communicate with each other. It is possible to achieve consensus between agents in system (1) by using the protocol, i.e., $\lim_{t\to\infty} \left\|x_i(t) - x_j(t)\right\| = 0$.

Assumption 1: (A, B) is stabilizable.

Assumption 2: The directed graph $\mathcal{G}$ is strongly connected.

Lemma 1 [35]:According to Assumption 2, the vector $\xi = [\xi_1, \xi_2, \ldots, \xi_N]^T$ whose elements are all positive. Furthermore, let $\Xi = \operatorname{diag}(\xi_1, \xi_2, \ldots, \xi_N)$. Then $\hat{L} = \left[\left(\Xi L + L^T \Xi\right)/2\right]$ is a symmetric matrix and $\sum_{j=1}^{N} \hat{L}_{ij} = \sum_{j=1}^{N} \hat{L}_{ji} = 0$ for all $i = 1, 2, \ldots, N$

Lemma 2 [36]: Based on Assumption 2, given a directed network with a Laplacian matrix L:

$$a(L) = \min_{x^T\xi=0, x\neq 0} \frac{x^T \hat{L} x}{x^T \Xi x} > 0, \tag{2}$$

where $\hat{L} = \left[\left(\Xi L + L^T \Xi\right)/2\right]$, $\xi = [\xi_1, \xi_2, \ldots, \xi_N]^T$, and $\Xi = \operatorname{diag}(\xi_1, \xi_2, \ldots, \xi_N)$, with $\xi_i > 0, i = 1, \ldots, N, \xi^T L = 0$, and $\sum_{i=1}^{N} \xi_i = 1$.

It should be noted that for a strongly linked graph, $a(L)$ is referred to as the general algebraic connectivity. The general algebraic connection of strongly connected graphs is known as $a(L)$. In an undirected graph, then $a(L) = \lambda_2(L)$.

III. A DISTRIBUTED DYNAMIC EVENT-TRIGGERED CONTROL APPROACH

In the spirit of [16], we first define the following combined measurement variable:

$$q_i(t) = \sum_{j=1}^{N} a_{ij} \left(x_j(t) - x_i(t)\right). \tag{3}$$

Following that, we provide a state feedback control mechanism for agent-i as described in [19] as

$$u_i(t) = Kq_i\left(t_k^i\right), \qquad t \in \left[t_k^i, t_{k+1}^i\right). \tag{4}$$

Agent-i only requires $q_i(t)$ at certain irregular periods, called as triggering times, $t_0^i, t_1^i, \ldots$, (or event times). Agent-i must use communication with its neighbors to calculate $q_i\left(t_k^i\right)$ at each trigger time instant. It is possible to determine the next triggering time instant t_{k+1}^i, if the current triggering time t_k^i is represented as follows:

$$\begin{cases} t_{k+1}^i = \inf\left\{t > t_k^i \; \|e_i(t)\|^2 - \delta_i \|q_i(t)\|^2 - \pi_i\eta_i(t) \geq 0\right\}, \\ \dot{\eta}_i(t) = -\beta_i\eta_i(t) + \theta_i\left(\delta_i \|q_i(t)\|^2 - \|e_i(t)\|^2\right), \eta_i(0) > 0, \end{cases} \tag{5}$$

where an internal dynamic state is defined by $\eta_i(t)$, while a measurement error must be defined by $e_i(t)$, $\beta_i > 0$, $\theta_i \geq \left\|PBB^T P\right\|$, $\pi_i > 0$, and $\delta_i = (\xi_i\sigma_i/\theta_i)$ with $0 < \sigma_i < 1$.
where

$$e_i(t) = q_i\left(t_k^i\right) - q_i(t). \tag{6}$$

Besides the controlled system's condition, the triggering function is likewise affected by the internal dynamic state $\eta_i(t)$. This type of triggering system is known as a dynamic event-triggered, and it was initially presented in [23] for a single controlled plant. Without the internal dynamic state, dynamic event-triggered is also known as static event-triggered.

The resulting closed-loop system, which is connected to the controller (4), is given by

$$\dot{x}_i = Ax_i + BK\left(e_i(t) + q_i(t)\right), \qquad i = 1, \ldots, N. \tag{7}$$

where $q(t) = \mathrm{col}\,(q_1(t), \ldots, q_N(t))$, and one has $q(t) = -(L \otimes I_n)\, x(t)$, according to (7), it follows that:

$$\dot{q}(t) = (I_N \otimes A - L \otimes BK)\, q(t)\text{-}(L \otimes BK)e(t). \tag{8}$$

K in the above equation has proven difficult. According to the literature [26], K may be computed by solving the following equation to get P.

$$PA + A^T P - \alpha' PBB^T P + I_n = 0, \tag{9}$$

where $\alpha' = 2\mu a(L) - \mu^2 \xi_M \|L\|^2 > 0$, with μ being a parameter that has to be calculated, a(L) provided in Lemma 2, and $\xi_M = \max_i (\xi_1, \ldots, \xi_N)$. This equation can be thought of as the special ARE with fixed parameters.

Theorem 1: According to Assumptions 1 and 2, let $K = \mu B^T P$, where $0 < \mu < \left(2a(L)/\left[\xi_M \|L\|^2\right]\right)$ and $P > 0$ is the solution to (9). Under the control protocol (4), the multi-agent system's agents reach consensus with t_{k+1}^i, determined by (3). As a result, any agent will not be able to perform Zeno behavior.

Proof: We take into account the subsequent candidate for the Lyapunov function:

$$W(t) = V(t) + \sum_{i=1}^{N} \eta_i(t), \tag{10}$$

where $V(t) = q^T(t)(\Xi \otimes P)q(t) \geq 0$, it is important to show that Zeno behavior does not occur in any agent in order to demonstrate that the triggering mechanism (5) could be utilized. If just agent-i exists, a Zeno behavior arises at some point in time T_0. Afterwards, there is $\lim_{k\to\infty} t_k^i = T_0$. Considering the limit property, it is concluded that for any any $\varepsilon_0 > 0$, there exists $N(\varepsilon_0)$ such that $t_k^i \in (T_0 - \varepsilon_0, T_0 + \varepsilon_0)$ for $\forall k \geq N(\varepsilon_0)$, which implies that $t_{N(\varepsilon_0)+1}^i - t_{N(\varepsilon_0)}^i < 2\varepsilon_0$. Note that $\sum_{i=1}^N \|q_i(t)\|^2 = \|q(t)\|^2 \leq (V(t)/[\xi_m \lambda_{\min}(P)])$ and $V(t) \leq W(t) \leq W(0)$. Then

$$\|q_i(t)\| \leq \|q(t)\| \leq \sqrt{\frac{W(0)}{\xi_{m\lambda_{\min}(P)}}} \doteq W_0. \tag{11}$$

Given that $\|e_i(t)\|$ is continuously piecewise differentiable in the interval $\left[t_k^i, t_{k+1}^i\right)$, the Dini derivative of $\|e_i(t)\|$ can be determined as follows:

$$\begin{aligned}
D^+ \|e_i(t)\| &\leq \frac{\left\|e_i^T\right\|}{\|e_i\|} \|\dot{e}_i\| \\
&= \left\| -\sum_{j=1}^{N} a_{ij} \left(\dot{x}_j(t) - \dot{x}_i(t)\right) \right\| \\
&= \| -Aq_i(t) \\
&\quad + \mu BB^T P \sum_{j=1}^{N} a_{ij} \left(q_i\left(t_k^i\right) - q_j\left(t_{k_j'}^j\right) \right) \| \\
&\leq \|A\| \, \|q_i(t)\| \\
&\quad + \mu \left\|BB^T P\right\| \sum_{j=1}^{N} a_{ij} \left(\left\|q_i\left(t_k^i\right)\right\| + \left\|q_j\left(t_{k_j'}^j\right)\right\| \right) \\
&\leq W_0 \left(\|A\| + \mu \left\|BB^T P\right\| (1 + |\mathcal{N}_i|) \right) \\
&\doteq \hat{W}_0,
\end{aligned} \tag{12}$$

where $k_j' = \arg\max_{k\in\mathbb{N}} \left\{ t_k^j \mid t_k^j \leq t \right\}$, and $D^+ \|e_i(t)\|$ derivatives on the right of $\|e_i(t)\|$ when $t = t_k^i$.

Due to the fact that an event is only triggered when $\|e_i(t)\|$ is reset to 0 and the triggering condition in (10) is met, one has $\|e_i(t)\| \geq (\delta_i \|q_i(t)\|^2 + \pi_i \eta_i)^{1/2} \geq \sqrt{\pi_i \eta_i}$ at t_k^{i-}, $k = 1, 2, \ldots$. Define $f\left(t^-\right) = \lim_{s\to t^-} f(s)$. Then

$$\left\| e_i\left(t_k^{i-}\right) \right\| \geq \sqrt{\pi_i \eta_i \left(t_k^{i-}\right)} = \sqrt{\pi_i \eta_i(0)} e^{-\frac{\beta_i + \theta_i \pi_i}{2} t_k^{i-}}. \tag{13}$$

It follows from (12) and (13) that:

$$t_{N(\varepsilon_0)+1}^i - t_{N(\varepsilon_0)}^i \geq \frac{1}{\hat{W}_0} \sqrt{\pi_i \eta_i(0)} e^{-\frac{\beta_i + \theta_i \pi_i}{2}} t_{N(\varepsilon_0)+1}^i. \tag{14}$$

Let $\varepsilon_0 > 0$ be a equation solution:

$$\frac{1}{\hat{W}_0} \sqrt{\pi_i \eta_i(0)} e^{-\frac{\beta_i + \theta_i \pi_i}{2} T_0} = 2\varepsilon_0 e^{\frac{\beta_i + \theta_i \pi_i}{2} \varepsilon_0}. \tag{15}$$

Then

$$\begin{aligned}
t_{N(\varepsilon_0)+1}^i - t_{N(\varepsilon_0)}^i &\geq \frac{1}{\hat{W}_0} \sqrt{\pi_i \eta_i(0)} e^{-\frac{\beta_i + \theta_i \pi_i}{2} (T_0 + \varepsilon_0)} \\
&= 2\varepsilon_0.
\end{aligned} \tag{16}$$

which contracts the fact that $t_{N(\varepsilon_0)+1}^i - t_{N(\varepsilon_0)}^i < 2\varepsilon_0$. Therefore, it follows that agent-i doesn't display Zeno behavior since the aforementioned presumption is false. The proof is finished.

IV. ADAPTIVE DYNAMIC PROGRAMMING

It is a difficult task to design a dynamic event-triggered control laws. Because ARE is difficult to solve, especially in high dimensions, earlier work relied on the solution of multiple matrix inequalities, but the presence of these inequalities may not be easily ensured. In adaptive dynamic programming, policy iteration converges faster than value iteration. In this section, a policy iterative approach is provided to approximate the algebraic Riccati problem's solution, and an ADP is created to solve the ARE equation (9). The policy iterative approach is paired with an online ADP algorithm. The developed method can approximate the control gain for each follower without relying on system matrix knowledge or the precise inertia, by utilizing all the finite data available, imposing an initial control policy on the agent at a limited time interval, collecting online measurements, and iterating by reusing the same online data.

A. ONLINE OFF-POLICY ALGORITHM

The continuous-time linear system (1), where A and B are constant matrices of appropriate size. Due to the existence of a constant matrix K has sufficient dimensions, so $A - BK$ is Hurwitz, (7) is considered as stable.

$$u = -Kx, \tag{17}$$

which reduces the performance index shown below to the minimum:

$$J(x_0; u) = \int_0^\infty \left(x^T Q x + u^T R u\right) dt. \tag{18}$$

where $Q = Q^T \geq 0, R = R^T > 0$ with $\left(A, Q^{1/2}\right)$ observable. By using (17) to (1), we can easily write (18) as follows:

$$J(x_0; u) = x_0^T P x_0, \tag{19}$$

where:

$$P = \int_0^\infty e^{(A-BK)^T t} \left(x^T Q x + K^T R K\right) e^{(A-BK)t} dt. \tag{20}$$

When the derivative of $X^T P X$ is taken along the solution of (1), there is only one positive definite solution P in the Lyapunov equation:

$$(A - BK)^T P + P(A - BK) + Q + K^T R K = 0. \tag{21}$$

When A and B are known exactly, the solution to an optimal control problem can be solved using classical optimal control theory by solving the following Riccati equation:

$$A^T P + PA + Q - PBR^{-1}B^T P = 0. \tag{22}$$

As a result, solving (22) can be difficult, particularly for high-dimensional matrices. Nevertheless, a number of effective methods have numerically approximated the solution to (34). One such technique is Kleinman's algorithm, which is detailed more below.

Theorem 2 [37]: Let $K_0 \in \mathbb{R}^{m \times n}$ be any stabilizing feedback gain matrix (i.e., $A - BK_0$ is Hurwitz), and repeat the following steps for $k = 0, 1, \ldots$

(1) Solve for the real symmetric positive definite solution P_k of the Lyapunov equation

$$A_k^T P_k + P_k A_k + Q + K_k^T R K_k = 0, \tag{23}$$

where:$A_k = A - BK_k$.

(2) The feedback gain matrix should be updated by:

$$K_{k+1} = R^{-1} B^T P_k. \tag{24}$$

Then, the following properties hold:

(1) $A - BK_k$ is Hurwitz;

(2) $P^* \leq P_{k+1} \leq P_k$;

(3) $\lim\limits_{k\to\infty} P_k = P^*$

Proof: Consider the Lyapunov equation (23) with $k = 0$. Since $A - BK_0$ is Hurwitz, by (20) we know P_0 is finite and positive definite. In addition, by (20) and (23) we have

$$P_0 - P_1 = \int_0^\infty e^{A_1^T \tau} (K_0 - K_1)^T R (K_0 - K_1) e^{A_1 \tau} d\tau \geq 0$$

Similarly, by (20) and (22) we obtain

$$P_1 - P^* = \int_0^\infty e^{A_1^T \tau} \left(K_1 - K^*\right)^T R \left(K_1 - K^*\right) e^{A_1 \tau} d\tau \geq 0$$

Therefore, we have $P^* \leq P_1 \leq P_0$. Since P^* is positive definite and P_0 is finite, P_1 must be finite and positive definite. This implies that $A - BK_1$ is Hurwitz. Repeating the above analysis for $k = 1, 2, \ldots$ proves Properties (1) and (2) in Theorem 2. Finally, since $\{P_k\}$ is a monotonically decreasing sequence and lower bounded by P^*, $\lim_{k\to\infty} P_k = P_\infty$ exists. By (23) and (24), $P = P_\infty$ satisfies (22), which has a unique solution. Therefore, $P_\infty = P^*$. The proof is thus complete.

Remark 1: When A and B are known, the Lyapunov equation (23) may be used to solve P_k and update K_k repeatedly. So, it is numerically approximated to find a solution for equation (22).

Theorem 2 will then be used to propose an offline policy iteration approach to address the optimal control problem.

According to (24) and (23), the method is offline and requires precise understanding of the dynamic. However, it is frequently difficult to construct a model of system dynamics or to get exact information about the dynamics of systems. In the spirit of Jiang et al. [43], an online ADP approach is proposed to handle this issue without the necessity for previous knowledge of system dynamics.

B. ADAPTIVE DYNAMIC PROGRAMMING BASED ON POLICY ITERATION

Based on the policy iteration algorithm presented above, we will describe an online ADP method that does not need A and B.

Along with the solutions of (1) by (24) and (23), one can obtain:

$$\begin{aligned} x(t+\Delta t)^T P_k x(t+\Delta t) - x(t)^T P_k x(t) &= -\int_t^{t+\Delta t} x^T Q_k x d\tau \\ &\quad + 2\int_t^{t+\Delta t} (u_0 + K_k x)^T R K_{k+1} x d\tau \end{aligned} \tag{25}$$

where $Q_k = Q + K_k^T R K_k$.

Remark 2: It is worth mentioning in Equation (25), that the system matrices can be replaced with the states and inputs measured online. As a consequence, Equation (25) may be used to calculate P_k and K_{k+1} without knowing the precise of A and B.

As a result, we use the Kronecker product to get (P_k, K_{k+1}) with unknown system matrices under a known stabilizing feedback gain matrix K_k:

$$x^T Q_k x = \left(x^T \otimes x^T\right) \mathrm{vec}(Q_k). \tag{26}$$

$$\begin{aligned} (u + K_k x)^T R K_{k+1} x = \Big[&\left(x^T \otimes x^T\right)\left(I_n \otimes K_k^T R\right) \\ &+ \left(x^T \otimes u_0^T\right)(I_n \otimes R)\Big] K_{k+1}, \end{aligned} \tag{27}$$

denote $\zeta_{xx} \in \mathbb{R}^{1\times n^2}$, $\varphi_{xx} \in \mathbb{R}^{1\times n^2}$ and $\varphi_{xu} \in \mathbb{R}^{1\times mn}$.

$$\begin{aligned} \zeta_{xx} = \Big[&x \otimes x|_{t_1}^{t_1+\delta t}, \quad x \otimes x|_{t_2}^{t_2+\delta t}, \\ &\ldots, \quad x \otimes x|_{t_i}^{t_i+\delta t}\Big]^T, \end{aligned} \tag{28}$$

$$\varphi_{xx} = \left[\int_{t_1}^{t_1+\delta t} x \otimes x d\tau, \quad \int_{t_2}^{t_2+\delta t} x \otimes x d\tau, \right.$$
$$\left. \ldots, \quad \int_{t_1}^{t_1+\delta t} x \otimes x d\tau \right]^T, \tag{29}$$
$$\varphi_{xu} = \left[\int_{t_1}^{t_1+\delta t} x \otimes u_0 d\tau, \quad \int_{t_2}^{t_2+\delta t} x \otimes u_0 d\tau, \right.$$
$$\left. \ldots, \quad \int_{t_l}^{t_l+\delta t} x \otimes u_0 d\tau \right]^T, \tag{30}$$

where $0 \leq t_1 \leq t_2 \leq \cdots \leq t_l$ then $\Phi_k \in \mathbb{R}^{l\times(n^2+mn)}$ and $\Psi_k \in \mathbb{R}^l$.

$$\Phi_k = \left[\zeta_{xx}, -2\varphi_{xx}\left(I_n \otimes K_k^T R\right) - 2\varphi_{xu}\left(I_n \otimes R\right)\right]. \tag{31}$$
$$\Psi_k = -\varphi_{xx}\,\mathrm{vec}\left(Q_k\right). \tag{32}$$

This is the form of Equation (25) when written as a linear equation:

$$\Phi_k \begin{bmatrix} \mathrm{vec}\left(P_k\right) \\ \mathrm{vec}\left(K_{k+1}\right) \end{bmatrix} = \Psi_k. \tag{33}$$

Lemma 3 [43]: It is possible to have a sufficiently large integer $l > 0$ such that:

$$\mathrm{rank}\left([\varphi_{xx}, \varphi_{xu}]\right) = \frac{n(n+1)}{2} + mn.$$

Lemma 4 [44]: Whenever k_0 is an initial stabilizing feedback control gain, and Lemma 3 holds, the sequences and obtained by solving (33) will respectively converge to optimal $\{P_k\}_{k=0}^{\infty}$ and $\{k_k\}_{k=0}^{\infty}$. Obtained by solving (33) will respectively converge to optimal P^* and K^*.

As a consequence, we can employ an online ADP to solve the optimal control problem without system dynamics.

Off-policy ADP algorithm

Step 1: Find K_0 such that $A - BK_0$ is Hurwitz. Let $k = 0$;

Step 2: Utilize $u_0 = -K_0 + e$ as the control input, e is the exploration noise. Compute, ζ_{xx}, φ_{xx} and φ_{xu} to satisfy the rank condition in Lemma 3;

Step 3: Solve for P_k and K_{k+1} from (33);

Step 4: Let $k = k + 1$ and repeat Step (3), until $|P_k - P_{k-1}| \leq \epsilon$, where the constant $\varepsilon > 0$ is a predefined small threshold;

Step 5: use $u = -K_k x$ as the approximate optimal control policy.

The above algorithm may be used to get the solution of (9) for P so that the control law K can be found.

V. SIMULITION

The problem of spacecraft formation flying is taken into account in this section to verify the theoretical approaches. The multi-agent system in this example consists of six agents, each of which represents a spacecraft that is in low Earth orbit. Several conversions can be done to translate the spacecraft formation flying problem into the consensus problem of a linear multi-agent system. The general linear dynamics of (1) apply to each agent in this case.

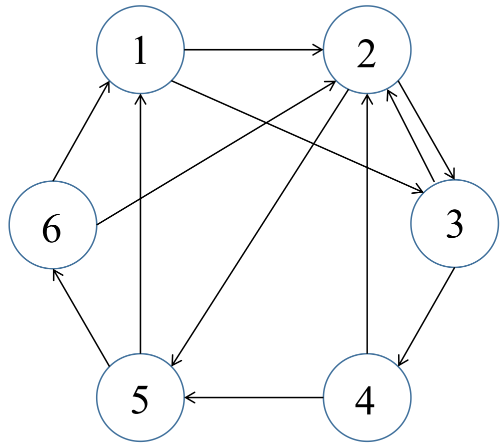

FIGURE 1. Graph $\mathcal{G}$ in the example.

$$A = \begin{bmatrix} 0 & I_3 \\ A_1 & A_2 \end{bmatrix}, \qquad B = \begin{bmatrix} 0 \\ I_3 \end{bmatrix},$$
$$A_1 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 3\omega_0^2 & 0 \\ 0 & 0 & -\omega_0^2 \end{bmatrix}, \qquad A_2 = \begin{bmatrix} 0 & 2\omega_0 & 0 \\ -2\omega_0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

The state is expressible as $x_i = \mathrm{col}\left(\bar{x}_i, \bar{y}_i, \bar{z}_i, v_i^x, v_i^y, v_i^z\right)$, where $(\bar{x}_i, \bar{y}_i, \bar{z}_i)$ denotes the distance in $X - Y - Z$ directions from the desired position; $\left(v_i^x, v_i^y, v_i^Z\right)$ is the velocity in the three directions; $\omega_0 = 0.001$ is the angular rate of the satellite. Assume that the directed communication graph shown in Figure 1 is used by the agents to communicate with one another.

The validity of Assumptions 1 and 2 can be verified, as a result, Theorem 1's requirements are met. It is possible to construct a triggered mechanism for the multi-agent system concerned based on Theorem 1. It is discovered via Lemma 2 that

$$\xi = \left[0.1250\ 0.1667\ 0.2083\ 0.0833\ 0.1667\ 0.25\right]$$

which satisfies $\xi^T L = 0$ and $\sum_{i=1}^{N} \xi_i = 1$, and other parameters are

$$a(L) = 0.6902, \qquad \mu = 0.0869 \leq \frac{2a(L)}{\xi_M \|L\|^2} = 0.2480$$

It is simple to solve for P in equation (9) utilizing the ADP provided in the previous chapter. Figure 2 demonstrates that P_k have reached their optimal values. As shown in the equation at the bottom of the next page.

Define the x-axis position error as $e_{ij} = \bar{x}_i - \bar{x}_j$, $\forall i \neq j$. Fig.3 and Fig.4 depict the evolution curves of agents' location error e_{ij} and velocity v_i^x using the event-triggered consensus procedure. It is demonstrated that all agents finally reach consensus on their positions and velocities on the x-axis. Furthermore, six agents' the triggering time

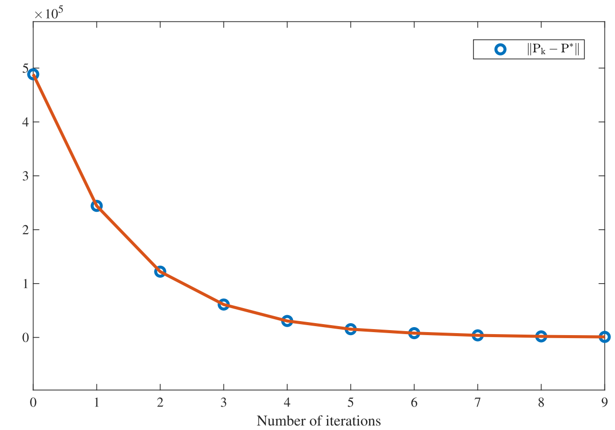

FIGURE 2. Triggering time instants for the agents.

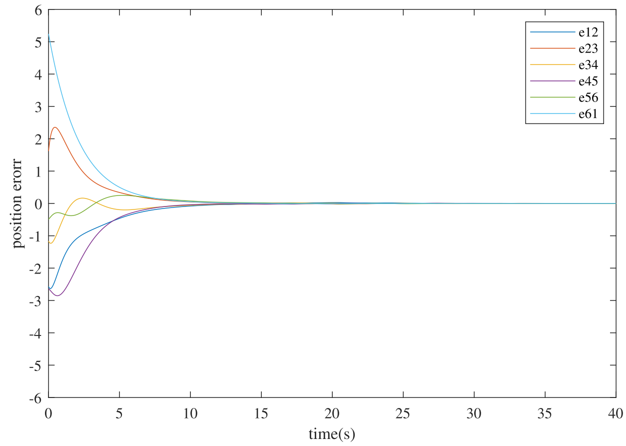

FIGURE 3. Evolution of agents' position error via the event-triggered control protocol.

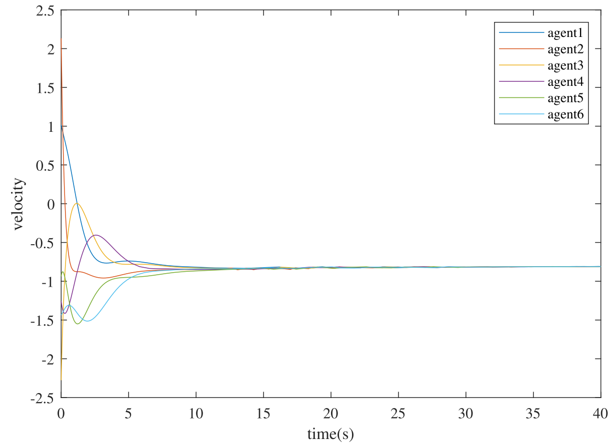

FIGURE 4. Evolution of agents' velocity error via the event-triggered control protocol.

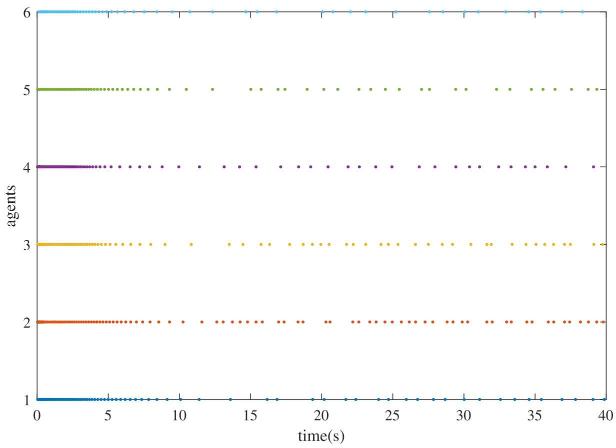

FIGURE 5. Triggering time instants for the agents.

instants are recorded, it is also clear to see in Fig. 5 how long the interval between events is for each agent. Furthermore, Figure 6 shows the trajectory of agent 1's combined measurement error and associated triggering threshold in triggering condition (5). The simulations, for comparison, are also built with the static event-triggering technique. Table 1 shows the triggering frequency using the static and dynamic triggering mechanisms, respectively. It can be seen that the dynamic triggered greatly lowers the triggering numbers.

$$
\begin{aligned}
P &= \begin{bmatrix}
2.8576 & 0.0000 & -0.0000 & 3.5829 & 0.0090 & -0.0000\\
0.0000 & 2.8576 & -0.0000 & -0.0090 & 3.5829 & -0.0000\\
-0.0000 & 0.0000 & 2.8576 & 0.0000 & -0.0000 & 3.5829\\
3.5829 & -0.0090 & -0.0000 & 10.2383 & -0.0000 & -0.0000\\
0.0090 & 3.5829 & -0.0000 & -0.0000 & 10.2384 & -0.0000\\
-0.0000 & 0.0000 & 3.5829 & 0.0000 & -0.0000 & 10.2383
\end{bmatrix}\\
K &= \mu B^T P\\
&= \begin{bmatrix}
0.3114 & -0.0008 & 0 & 0.8897 & 0 & 0\\
0.0008 & 0.3114 & 0 & 0 & 0.8897 & 0\\
0 & 0 & 0.3114 & 0 & 0 & 0.8897
\end{bmatrix}\\
\delta &= \begin{bmatrix} 1.1 & 1.4 & 1.8 & 0.7 & 1.4 & 2.1 \end{bmatrix} \times 10^{-3}\\
\theta_i &= \left\| PBB^T P \right\| = 117.6614, \quad \beta_i = 0.004, \pi_i = 0.003.
\end{aligned}
$$

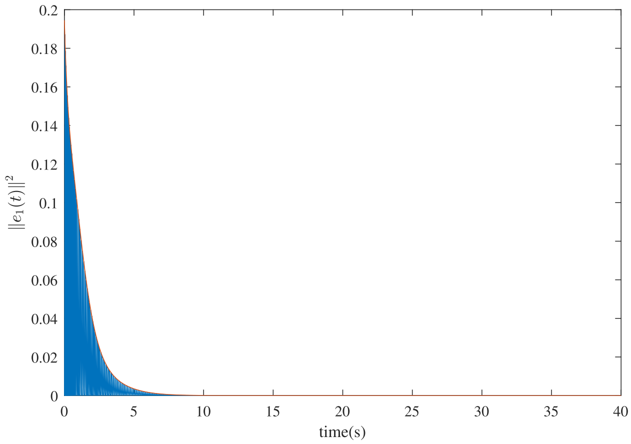

FIGURE 6. Triggering time instants for the agents.

TABLE 1. Units for magnetic properties.

agent	1	2	3	4	5	6
dynamic	81	98	86	80	71	61
static	453	372	356	540	373	346

VI. CONCLUSION

This paper describes how to design control laws for dynamic event-triggered mechanisms. Designing a dynamic event-triggered control law under a dynamic triggering mechanism is a difficult problem to solve. The multi-agent systems under consideration have directed communication graph and generic linear dynamics. This issue requires precise dynamics, which is difficult since most systems in practice are too complicated, and the resulting dynamics may be inaccurate. We designed model-free adaptive dynamic programming, which perfectly solves this problem, to address these two issues. The findings demonstrate that the event-triggered control protocol with this laws allows exponential consensus among all agents and that none of the agents exhibit Zeno behavior. But the dynamic event triggering mechanism in this article relies on the ARE equation with fixed parameters. Future research will focus on the design of flexible ARE equations for dynamic event triggering mechanisms to design control laws.

REFERENCES

[1] T. Chu, J. Wang, L. Codecà, and Z. Li, "Multi-agent deep reinforcement learning for large-scale traffic signal control," *IEEE Trans. Intell. Transp. Syst.*, vol. 21, no. 3, pp. 1086–1095, Mar. 2020.

[2] H. Liu, Q. Meng, F. Peng, and F. L. Lewis, "Heterogeneous formation control of multiple UAVs with limited-input leader via reinforcement learning," *Neurocomputing*, vol. 412, pp. 63–71, Oct. 2020.

[3] X. Zhang, Z. Cheng, J. Ma, S. Huang, F. L. Lewis, and T. H. Lee, "Semi-definite relaxation-based ADMM for cooperative planning and control of connected autonomous vehicles," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 7, pp. 9240–9251, Jul. 2022.

[4] V. Borkar and P. P. Varaiya, "Asymptotic agreement in distributed estimation," *IEEE Trans. Autom. Control*, vol. AC-27, no. 3, pp. 650–655, Jun. 1982.

[5] J. N. Tsitsiklis, "Problems in decentralized decision making and computation," Massachusetts Inst. Technol. Lab. Inf. Decis. Syst., Cambridge, MA, USA, Tech. Rep., 1984.

[6] X.-M. Zhang and Q.-L. Han, "Network-based $\mathcal{H}_\infty$ filtering using a logic jumping-like trigger," *Automatica*, vol. 49, no. 5, pp. 1428–1435, May 2013.

[7] B.-L. Zhang, Q.-L. Han, and X.-M. Zhang, "Recent advances in vibration control of offshore platforms," *Nonlinear Dyn.*, vol. 89, no. 2, pp. 755–771, Apr. 2017.

[8] G. Xie, H. Liu, L. Wang, and Y. Jia, "Consensus in networked multi-agent systems via sampled control: Fixed topology case," in *Proc. Amer. Control Conf.*, 2009, pp. 3902–3907.

[9] X.-M. Zhang, Q.-L. Han, and B.-L. Zhang, "An overview and deep investigation on sampled-data-based event-triggered control and filtering for networked systems," *IEEE Trans. Ind. Informat.*, vol. 13, no. 1, pp. 4–16, Feb. 2017.

[10] X.-M. Zhang, Q.-L. Han, and X. Yu, "Survey on recent advances in networked control systems," *IEEE Trans. Ind. Informat.*, vol. 12, no. 5, pp. 1740–1752, Oct. 2016.

[11] X. Ge, Q.-L. Han, D. Ding, X.-M. Zhang, and B. Ning, "A survey on recent advances in distributed sampled-data cooperative control of multi-agent systems," *Neurocomputing*, vol. 275, pp. 1684–1701, Jan. 2017.

[12] K. J. Åström and B. Bernhardsson, "Comparison of periodic and event based sampling for first-order stochastic systems," *IFAC Proc. Volumes*, vol. 32, no. 2, pp. 5006–5011, 1999.

[13] P. Tabuada, "Event-triggered real-time scheduling of stabilizing control tasks," *IEEE Trans. Autom. Control*, vol. 52, no. 9, pp. 1680–1685, Sep. 2007.

[14] Z. Zuo, Q.-L. Han, B. Ning, X. Ge, and X.-M. Zhang, "An overview of recent advances in fixed-time cooperative control of multiagent systems," *IEEE Trans. Ind. Informat.*, vol. 14, no. 6, pp. 2322–2334, Jun. 2018.

[15] D. Dimarogonas, E. Frazzoli, and K. Johansson, "Distributed event-triggered control for multi-agent systems," *IEEE Trans. Autom. Control*, vol. 57, no. 5, pp. 1291–1297, May 2020.

[16] Y. Fan, G. Feng, Y. Wang, and C. Song, "Distributed event-triggered control of multi-agent systems with combinational measurements," *Automatica*, vol. 49, no. 2, pp. 671–675, Feb. 2013.

[17] G. S. Seyboth, D. V. Dimarogonas, and K. H. Johansson, "Event-based broadcasting for multi-agent average consensus," *Automatica*, vol. 49, no. 1, pp. 245–252, Jan. 2013.

[18] H. Li, X. Liao, T. Huang, and W. Zhu, "Event-triggering sampling based leader-following consensus in second-order multi-agent systems," *IEEE Trans. Autom. Control*, vol. 60, no. 7, pp. 1998–2003, Jul. 2015.

[19] W. Hu, L. Liu, and G. Feng, "Consensus of linear multi-agent systems by distributed event-triggered strategy," *IEEE Trans. Cybern.*, vol. 46, no. 1, pp. 148–157, Jan. 2016.

[20] W. Zhu, Z.-P. Jiang, and G. Feng, "Event-based consensus of multi-agent systems with general linear models," *Automatica*, vol. 50, no. 2, pp. 552–558, Feb. 2014.

[21] H. Li, G. Chen, and L. Xiao, "Event-triggered nonlinear consensus in directed multi-agent systems with combinational state measurements," *Int. J. Syst. Sci.*, vol. 47, no. 14, pp. 3364–3377, Oct. 2016.

[22] X. Liu, C. Du, H. Liu, and P. Lu, "Distributed event-triggered consensus control with fully continuous communication free for general linear multi-agent systems under directed graph," *Int. J. Robust Nonlinear Control*, vol. 28, no. 1, pp. 132–143, 2018.

[23] A. Girard, "Dynamic triggering mechanisms for event-triggered control," *IEEE Trans. Autom. Control*, vol. 60, no. 7, pp. 1992–1997, Jul. 2015.

[24] W. Hu and C. Yang, "Consensus of linear multi-agent systems by distributed dynamic event-triggered control," in *Proc. Int. Workshop Complex Syst. Netw. (IWCSN)*, Dec. 2017, pp. 284–289.

[25] Q. Li, B. Shen, Z. Wang, T. Huang, and J. Luo, "Synchronization control for a class of discrete time-delay complex dynamical networks: A dynamic event-triggered approach," *IEEE Trans. Cybern.*, vol. 49, no. 5, pp. 1979–1986, May 2019.

[26] W. Hu, C. Yang, T. Huang, and W. Gui, "A distributed dynamic event-triggered control approach to consensus of linear multiagent systems with directed networks," *IEEE Trans. Cybern.*, vol. 50, no. 2, pp. 869–874, Feb. 2020.

[27] D. Li and J. Dong, "Robust control for a class of nonlinear systems with input constraints based on actor-critic learning," *Int. J. Robust Nonlinear Control*, vol. 2022, pp. 1–15, Jun. 2022.

[28] D. V. Prokhorov and D. C. Wunsch, "Adaptive critic designs," *IEEE Trans. Neural Netw.*, vol. 8, no. 5, pp. 997–1007, Sep. 1997.

[29] D. Molina, G. K. Venayagamoorthy, J. Liang, and R. G. Harley, "Intelligent local area signals based damping of power system oscillations using virtual generators and approximate dynamic programming," *IEEE Trans. Smart Grid*, vol. 4, no. 1, pp. 498–508, Mar. 2013.

[30] D. P. Bertsekas and J. N. Tsitsiklis, "Neuro-dynamic programming: An overview," in *Proc. 34th IEEE Conf. Decis. Control*, vol. 1, Dec. 1995, pp. 560–564.

[31] F. L. Lewis, D. Vrabie, and K. G. Vamvoudakis, "Reinforcement learning and feedback control: Using natural decision methods to design optimal adaptive controllers," *IEEE Control Syst. Mag.*, vol. 32, no. 6, pp. 76–105, Dec. 2012.

[32] J. J. Murray, C. J. Cox, G. G. Lendaris, and R. Saeks, "Adaptive dynamic programming," *IEEE Trans. Syst., Man, Cybern., C, Appl. Rev.*, vol. 32, no. 2, pp. 140–153, Oct. 2002.

[33] R. Song, W. Xiao, H. Zhang, and C. Sun, "Adaptive dynamic programming for a class of complex-valued nonlinear systems," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 25, no. 9, pp. 1733–1739, Sep. 2014.

[34] D. Liu and Q. Wei, "Policy iteration adaptive dynamic programming algorithm for discrete-time nonlinear systems," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 25, no. 3, pp. 621–634, Mar. 2014.

[35] W. Lu and T. Chen, "New approach to synchronization analysis of linearly coupled ordinary differential systems," *Phys. D, Nonlinear Phenomena*, vol. 213, no. 2, pp. 214–230, 2006.

[36] W. Yu, G. Chen, M. Cao, and J. Kurths, "Second-order consensus for multiagent systems with directed topologies and nonlinear dynamics," *IEEE Trans. Syst., Man, Cybern., B, Cybern.*, vol. 40, no. 3, pp. 881–891, Jun. 2010.

[37] D. L. Kleinman, "On an iterative technique for Riccati equation computations," *IEEE Trans. Autom. Control*, vol. AC-13, no. 1, pp. 114–115, Feb. 1968.

[38] J. Wang, Z. Zhang, B. Tian, and Q. Zong, "Event-based robust optimal consensus control for nonlinear multiagent system with local adaptive dynamic programming," *IEEE Trans. Neural Netw. Learn. Syst.*, early access, Jun. 27, 2022, doi: 10.1109/TNNLS.2022.3180054.

[39] W. Zhao and H. Zhang, "Distributed optimal coordination control for nonlinear multi-agent systems using event-triggered adaptive dynamic programming method," *ISA Trans.*, vol. 91, pp. 184–195, Aug. 2019.

[40] Z. Shi and C. Zhou, "Distributed optimal consensus control for nonlinear multi-agent systems with input saturation based on event-triggered adaptive dynamic programming method," *Int. J. Control*, vol. 95, no. 2, pp. 282–294, Feb. 2022.

[41] T. Wu, J. Hu, and D. Chen, "Non-fragile consensus control for nonlinear multi-agent systems with uniform quantizations and deception attacks via output feedback approach," *Nonlinear Dyn.*, vol. 96, no. 1, pp. 243–255, Apr. 2019.

[42] J. Hu, C. Jia, H. Yu, and H. Liu, "Dynamic event-triggered state estimation for nonlinear coupled output complex networks subject to innovation constraints," *IEEE/CAA J. Autom. Sinica*, vol. 9, no. 5, pp. 941–944, May 2022.

[43] Y. Jiang and Z.-P. Jiang, "Global adaptive dynamic programming for continuous-time nonlinear systems," *IEEE Trans. Autom. Control*, vol. 60, no. 11, pp. 2917–2929, Nov. 2015.

[44] W.-X. Jing and C. R. McInnes, "Memorised quasi-time–fuel-optimal feedback control of perturbed double integrator," *Automatica*, vol. 38, no. 8, pp. 1389–1396, Aug. 2002.

QI ZHANG received the bachelor's degree from Shijiazhuang Tiedao University, Hebei, China, in 2018. He is currently pursuing the master's degree with the Changchun University of Science and Technology, Changchun, China.

His research interests include multi-agent systems, event-triggered, adaptive dynamic programming, and reinforcement learning.

YANG YANG (Member, IEEE) received the M.S. and Ph.D. degrees from the Changchun University of Science and Technology, Changchun, China, in 2005 and 2013, respectively.

He was a Visiting Scholar at the University of Notre Dame, USA, in 2011. He is currently a Professor with the Changchun University of Science and Technology. His research interests include cyber physical systems, intelligent algorithms, and data-driven control.

XIAORAN XIE received the bachelor's degree from ShanDong JiaoTong University, Shandong, China, in 2020. She is currently pursuing the master's degree with the Changchun University of Science and Technology, Changchun, China.

Her research interests include multi-agent networks and distributed algorithm.

CHUNMING XU is working with Changchun Shikai Science and Technology Industry Company Ltd., China. His research interests include the Internet of Things and emergencies.

HAN YANG is working with the Beijing Engineering Research Center of Emergency Survival Security, China. His research interests include the Internet of Things and emergencies.