2018 Annual American Control Conference (ACC)
June 27–29, 2018. Wisconsin Center, Milwaukee, USA

Model-Free Event-Triggered Containment Control of Multi-Agent Systems

Yongliang Yang[1], Hamidreza Modares[2], Kyriakos G. Vamvoudakis,[3], Yixin Yin[1], Donald C.Wunsch II[2]

***Abstract*— This paper presents a model-free distributed event-triggered containment control scheme for linear multi-agent systems. The proposed event-triggered scheme guarantees asymptotic stability of the equilibrium point of the containment error as well the avoidance of the Zeno behavior. To relax the requirement of complete knowledge of the dynamics, we combine an off-policy reinforcement learning algorithm in an actor critic structure with the event-trigger control mechanism to obtain the feedback gain of the distributed containment control protocol. A simulation experiment is conducted to verify the effectiveness of the approach.**

***Index Terms*— Multi-agent systems, containment control, event-triggered, off-policy reinforcement learning, actor-critic.**

I. INTRODUCTION

Distributed control of multi-agent systems (MASs) has attracted great amount of attention from multiple fields, such as vehicle swarms [1], social networks [2], and so on. Distributed control protocols for MASs are designed based on the exchange of local information among neighboring agents, and the interaction among agents is dictated by a communication graph [3]. Examples of distributed MASs are flocking control [4], formation control [5], graphical games [6]–[8], synchronization control [9], [10] and so on. We will focus on the containment control problem of MASs, for which the objective is to drive the state of all followers into the convex hull spanned by leaders' states [11].

Related work

Containment control of MASs has practical uses in real-world applications. For example, when a group of robots performs the task of fire fighting in a forest, it is necessary to make all the robots avoid the obstacles and arrive at the designated area to accomplish their missions. The leader robots which have more situation awareness capabilities can guide other follower robots to accomplish the mission. Since the interior point in the convex hull spanned by leaders, i.e., the desired trajectory for all followers, is not unique, the results from leader-follower distributed consensus control cannot be applied to the containment problem directly. Necessary and sufficient conditions that guarantee the solution of the containment control for first- and second-order MASs with stationary and dynamic leaders have been investigated in [12]. Recently, internal model principle [13], sliding model approach [14] and observer-based design [15] have been applied to solve the containment problem. However, these results require constant interaction among agents, which may not be practical when the communication resources are limited.

The event-triggered design, is originally proposed to stabilize single-agent systems while decreasing the communication burden [16], [17]. Such work has been recently extended to develop the distributed consensus control protocol for MASs [18]–[21]. Up to now, the state-dependent sampling results in the dependency of each agent's control update on her neighbors' events, which might not reduce the communication and computation load efficiently. Therefore, a novel event-triggered mechanism is developed to guarantee such independence.

Another issue in existing approaches for both time-triggered and event-triggered containment control protocol design methods is that complete knowledge of agents' dynamics is required. Reinforcement learning (RL) has been used effectively to solve optimal regulation and tracking problems, as well as differential games and H_∞ problems for single and MASs systems [22], [23], [24], [6], [25], [9]. In this paper, off-policy RL in an actor-critic structure is developed to obtain the event-triggered containment control in a model-free manner.

Contributions

The contributions are two-fold. First, both time-triggered and event-triggered distributed containment control protocols are developed. The proposed event-triggered scheme is based on the sampling of the containment error, which guarantees that the event-triggering instant for each agent is independent of the other agents' states. Second, to obviate the requirement of complete knowledge of agents' dynamics, an off-policy RL is adopted to obtain the distributed event-triggered containment control protocol in a model-free manner.

Structure: The remainder of this paper is structured as follows. Section II formulates the problem. A time-triggered distributed protocol to solve the containment problem is given in Section III. To avoid the continuous interaction amongst the agents, a novel event-triggered containment control framework along with Zeno-free behavior guarantees, is proposed in Section IV. To obviate the requirement of complete information of the agent dynamics, Section V

This work was supported in part by the Mary K. Finley Missouri Endowment, the Missouri S&T Intelligent Systems Center, the National Science Foundation and the National Natural Science Foundation of China (NSFC Grant No. 61333002) and by NATO under grant No. SPS G5176.

[1]Y. Yang and Y. Yin are with the School of Automation & Electrical Engineering, University of Science and Technology Beijing, Beijing 100083, China. y.yang.2016@ieee.org; yyx@ies.ustb.edu.cn

[2]H. Modares and D. C. Wunsch II are with the Department of Electrical & Computer Engineering, Missouri University of Science & Technology, Rolla, MO 65409, USA. modares@mst.edu; dwunsch@mst.edu

[3]K. G. Vamvoudakis is with the Kevin T. Crofton Department of Aerospace and Ocean Engineering, Virginia Tech, Blacksburg, VA 24061-0203, USA. kyriakos@vt.edu

uses an off-policy RL algorithm to learn the feedback gain of the event-triggered containment control. To validate the proposed protocol design, a simulation example is carried out in Section VI. Finally, Section VII concludes and talks about future directions.

Notation and Background: The background information and assumptions for graphs used here is standard.

$\mathcal{G} \triangleq$ Graph of all agents, $\mathcal{G} = \{1, ..., M+N\}$
$\mathcal{R} \triangleq$ subgraph of leaders, $\mathcal{R} = \{1, ..., M\}$
$\mathcal{F} \triangleq$ subgraph of followers, $\mathcal{F} = \{M+1, ..., M+N\}$
$\mathcal{A} \triangleq$ adjacency matrix of subgraph $\mathcal{F}$
$a_{ij} \triangleq$ entry of adjacency matrix $\mathcal{A}$
$d_i \triangleq$ in-degree of node i
$D \triangleq$ in-degree matrix of subgraph $\mathcal{F}$ ($D = \mathrm{diag}\{d_1, ..., d_N\}$)
$\mathcal{N}_i \triangleq$ set of neighbors of node i
$n_i \triangleq$ the number of agent i's neighbors
$\mathcal{L} \triangleq$ Laplacian matrix of subgraph $\mathcal{F}$ ($\mathcal{L} = D - \mathcal{A}$)
$G_\rho \triangleq$ pinning matrix of ρ-th leader to all followers ($G_\rho = \mathrm{diag}\{g_1^\rho, \cdots, g_N^\rho\}$), $G = \sum_{\rho\in\mathcal{R}} G_\rho$
$g_i^\rho \triangleq$ pinning gain from ρ-th leader to i-th follower
$g_i \triangleq$ i-th element on the diagonal of G ($g_i = \sum_{\rho\in\mathcal{R}} g_i^\rho$)
$\lambda_i \triangleq$ i-th eigenvalue of matrix $\mathcal{L}+G$ ($\lambda_1 \leqslant \cdots \leqslant \lambda_N$)

Assumption 1. The subgraph $\mathcal{F}$ is undirected. For each follower, there exists at least one leader that has a directed path to that follower. □

Lemma 1. *[26] Under Assumption 1, all the eigenvalues of the matrix $\mathcal{L}+G$, $\lambda_1, \cdots, \lambda_N$, are positive.* □

II. Problem Formulation

Consider the multi-agent system with N followers,

$$\dot{x}_i = Ax_i + Bu_i, \ i \in \mathcal{F} \tag{1}$$

where $x_i \in \mathbb{R}^n$ is the state and $u_i \in \mathbb{R}^m$ is control of follower i.

The leaders dynamics are modeled as,

$$\dot{x}_\rho = Ax_\rho, \ \rho \in \mathcal{R} \tag{2}$$

where $x_\rho \in \mathbb{R}^n$ is the leader's state.

Assumption 2. The pair (A, B) is assumed to be stabilizable. □

The following definitions are adopted from [27].

Definition 1. *(Distance)* Let $x \in \mathbb{R}^n$ and $\mathcal{C} \subseteq \mathbb{R}^n$. Then, the distance from x to the set $\mathcal{C}$ is defined as

$$\mathrm{dist}\,(x, \mathcal{C}) = \inf_{y\in\mathcal{C}} \|x - y\|\,. \qquad \square$$

Definition 2. *(Convex Hull)* A set $\mathcal{C} \subseteq R^n$ is convex if $(1-\lambda)\,x + \lambda y \in \mathcal{C}$, $\forall x, y \in \mathcal{C}$ and $\lambda \in [0,1]$. The convex hull $\mathrm{Co}(\mathcal{X})$ of a finite set of q points $\mathcal{X} = \{x_1, x_2, \cdots, x_q\}$ is the minimal convex set containing all points in $\mathcal{X}$. That is, $\mathrm{Co}\,(\mathcal{X}) = \left\{ \sum_{i=1}^{q} \alpha_i x_i \,\middle|\, x_i \in \mathcal{X}, \alpha_i \in \mathbb{R}, \sum_{i=1}^{q} \alpha_i = 1 \right\}$. □

The containment control problem of MAS with leaders in (2) and followers in (1) can now be formulated as follows.

Problem 1. Design the distributed control protocol u_i such that $\lim_{t\to\infty} \mathrm{dist}\left(x_i\,(t), \mathrm{Co}\,\{x_\rho\,(t)\}_{\rho\in\mathcal{R}}\right) = 0, \forall i \in \mathcal{F}$, i.e., all the followers are synchronized to the convex hull spanned by the leaders. □

The local interaction between agent i and her neighbors can be expressed as,

$$\delta_i = \sum_{j\in\mathcal{N}_i} a_{ij}\,(x_i - x_j) + \sum_{\rho\in\mathcal{R}} g_i^\rho\,(x_i - x_\rho). \tag{3}$$

and in a compact form as,

$$\delta = Hx - \sum_{\rho=1}^{M} H_\rho \bar{x}_\rho \tag{4}$$

with,

$$h_\rho = \frac{\mathcal{L}}{M} + G_\rho, H_\rho = h_\rho \otimes I_n,$$
$$h = \mathcal{L} + G, H = h \otimes I_n,$$

where $\bar{x}_\rho = 1_N \otimes x_\rho$, $x = \left[\begin{array}{ccc} x_1^{\mathrm{T}} & \cdots & x_N^{\mathrm{T}} \end{array}\right]^{\mathrm{T}}$. Then, the dynamics of δ in (4) are,

$$\begin{aligned} \dot{\delta} &= H\dot{x} - \sum_{\rho=1}^{M} H_\rho \dot{\bar{x}}_\rho \\ &= \bar{A}\left(Hx - \sum_{\rho=1}^{M} H_\rho \bar{x}_\rho\right) + H\bar{B}u \\ &= \bar{A}\delta + H\bar{B}u. \end{aligned} \tag{5}$$

where $\bar{A} = I_N \otimes A$ and $\bar{B} = I_N \otimes B$.

According to [13], it is guaranteed that the containment control problem is solved if (5) has an asymptotically stable equilibrium point. Therefore, δ in (4) is referred to as global containment error, and δ_i in (3) is referred to as local containment error.

III. Time-triggered Containment Control Design

Considering the local information δ_i in (3), the distributed containment control can be written as,

$$u_i = cK\delta_i, \tag{6}$$

where $c \in \mathbb{R}^+$ is the coupling gain and $K \in \mathbb{R}^{m\times n}$ is the feedback gain matrix to be designed later.

Using (6) in (4) yields,

$$\dot{\delta} = \left[(I_N \otimes A) + (ch \otimes BK)\right]\delta. \tag{7}$$

The following results provide the proper design of continuous-time containment control in (6).

Lemma 2. *Under Assumptions 1 and 2, suppose that there exist matrices $Q \succ 0 \in \mathbb{R}^{n\times n}$ and $R \succ 0 \in \mathbb{R}^{m\times m}$. Given that, the feedback and the coupling gains are selected as*

$$K = -R^{-1}B^{\mathrm{T}}P, \tag{8}$$

$$c \geqslant \frac{1}{2\min\limits_{i=1,\cdots,N}\lambda_i}, \tag{9}$$

with $P \succ 0$ satisfying the algebraic Riccati equation (ARE),

$$A^{\mathrm{T}}P + PA - PBR^{-1}B^{\mathrm{T}}P + Q = 0, \tag{10}$$

then, the equilibrium point (i.e. origin) of the closed-loop system with state δ is asymptotically stable. □

Proof. The proof is an extension of [28, Theorem 1] and is omitted here due to page limitation. □

Remark 1. The distributed control protocol in (6) requires continuous interaction between each agent and her neighbors, and is referred to as time-triggered containment control in contrast to the event-triggered containment control discussed in the next Section. □

IV. Event-Triggered Distributed Containment Control Design

A. Event-Triggered Scheme Design

In this section, an event-triggered containment control is developed. The aperiodic sampling of local containment error δ_i is updated only when an event is triggered. That is, the local containment error remains constant between two successive events

$$\hat{\delta}_i(t) = \delta_i\left(t_k^i\right),\ t \in \left[t_k^i, t_{k+1}^i\right) \tag{11}$$

where $\{t_k^i\}_{k=0}^{\infty}$ is a monotonically increasing sequence of sampling instants, satisfying $\lim_{k\to\infty} t_k^i = \infty$, with $t_0^i = t_0$ $\forall i \in \mathcal{F}$. Then, the event-triggered containment control can be designed as,

$$u_i(t) = cK\hat{\delta}_i(t) = cK\delta_i\left(t_k^i\right),\ t \in \left[t_k^i, t_{k+1}^i\right). \tag{12}$$

From (12), it can be seen that for each follower, the distributed control is updated only at the triggering time t_k^i. Let $e_i(t)$ denote the sampling error between $\hat{\delta}_i(t)$ and $\delta_i(t)$, i.e.,

$$e_i(t) = \hat{\delta}_i(t) - \delta_i(t).$$

Then, the event-triggered distributed control (12) can be equivalently expressed as,

$$u_i(t) = cK\left[e_i(t) + \delta_i(t)\right]. \tag{13}$$

Note that although the control policy (13) is continuous, it is measurable and locally essentially bounded. Then, the Filippov solutions for each follower with event-triggered control (13) exists [29], [30]. Therefore, using (13) in (1) and using differential inclusions and nosnsmooth analysis [29], one has

$$\dot{x}_i(t) \in^{a.e.} \mathcal{K}\left[Ax_i(t) + cBK\left[e_i(t) + \delta_i(t)\right]\right]. \tag{14}$$

Writing a compact form of (14) for all the followers yields,

$$\dot{x}(t) \in^{a.e.} \mathcal{K}\left[(I_N \otimes A)x(t) + (I_N \otimes cBK)(e+\delta)\right], \tag{15}$$

where e is a compact form of all the vectors in (13).

Similarly, the dynamics of the leaders are given by,

$$\dot{\bar{x}}_0(t) = (I_N \otimes A)\bar{x}_0(t). \tag{16}$$

Substituting the event-triggered control from (13) in (15) and (16) gives,

$$\dot{\delta}(t) \in^{a.e.} \mathcal{K}\left[(I_N \otimes A + ch \otimes BK)\delta + (ch \otimes BK)e\right]. \tag{17}$$

Based on $\hat{\delta}_i(t)$ in (11), the event-triggered containment control protocol (12) can be designed by the following theorem.

Theorem 1. *Under Assumption 1 and 2, let the feedback gain in (12) be designed as*

$$K = -B^{\mathrm{T}}P, \tag{18}$$

where $P \succ 0$ is the unique solution to the following ARE

$$A^{\mathrm{T}}P + PA - PBB^{\mathrm{T}}P + \beta I_n = 0, \tag{19}$$

where we have substituted $Q := \beta I_n$ in (10) *and the coupling gain c in (12) is selected such that,*

$$\lambda_N \geqslant \lambda_i \geqslant \lambda_1 \geqslant \frac{1}{2c} > 0, \forall i \in \{1,\cdots,N\}. \tag{20}$$

Furthermore, let β from (19) satisfy,

$$\beta \geqslant \lambda_N \geqslant \lambda_i \geqslant \lambda_1 > 0, \forall i \in \{1,\cdots,N\}. \tag{21}$$

where $\lambda_i,\ i = 1,...,N$ are the eigenvalues of the matrix h. Then, $\delta_i \to 0, \forall i \in \mathcal{F}$, as $t \to \infty$ as long as the event-triggering condition,

$$\|e_i\| \geqslant \pi_i\|\delta_i\|, \tag{22}$$

is satisfied, with $\pi_i = \sqrt{\sigma_i\frac{r}{c\gamma}(1 - c\gamma r)}$ with $\gamma = \left\|PBB^{\mathrm{T}}P\right\|$, $r < \frac{1}{c\gamma}$ and $\sigma_i \in (0,1)$. □

Proof. Consider the following Lyapunov equation,

$$V(\delta) = \frac{1}{2}\delta^{\mathrm{T}}(h \otimes P)\delta.$$

Taking the time derivative of V and substituting (17) yields,

$$\dot{V} = \mathcal{K}\Bigg[\underbrace{\frac{1}{2}\delta^{\mathrm{T}}\left[h \otimes \left(A^{\mathrm{T}}P + PA\right)\right]\delta}_{V_1} + \underbrace{\delta^{\mathrm{T}}\left[ch^2 \otimes PBK\right]\delta}_{V_2} + \underbrace{\delta^{\mathrm{T}}\left[ch^2 \otimes PBK\right]e}_{V_3}\Bigg]. \tag{23}$$

Based on Assumption 1 and Lemma 1, there exists an orthogonal matrix $U^{\mathrm{T}} = U^{-1}$ such that

$$U^{\mathrm{T}}hU = \Lambda, \tag{24}$$

$$U\Lambda U^{\mathrm{T}} = h, \tag{25}$$

where $\Lambda = \operatorname{diag}\left(\begin{bmatrix} \lambda_1 & \cdots & \lambda_N \end{bmatrix}\right)$ is the diagonal matrix consisting of the eigenvalues of h.

The following transformation can be used,

$$\eta = \left(U^{\mathrm{T}} \otimes I_n\right) \delta, \tag{26}$$

$$\varepsilon = \left(U^{\mathrm{T}} \otimes I_n\right) e. \tag{27}$$

Combining the matrix transformation (24), (25) and variable transformation (26), (27), V_1, V_2 and V_3 in (23) can be written as,

$$V_1 = \frac{1}{2}\eta^{\mathrm{T}} \left[\Lambda \otimes \left(A^{\mathrm{T}}P + PA\right)\right] \eta,$$
$$V_2 = \eta^{\mathrm{T}} \left[\Lambda^2 \otimes cPBK\right] \eta,$$
$$V_3 = \eta^{\mathrm{T}} \left[\Lambda^2 \otimes cPBK\right] \varepsilon.$$

First, note that,

$$\begin{aligned} V_1 + V_2\ le\ & \frac{1}{2}\sum_{i=1}^{N} \left[\lambda_i \left(A^{\mathrm{T}}P + PA\right) + 2c\lambda_i^2 PBK\right] \|\eta_i\|^2 \\ \leqslant & \frac{1}{2}\sum_{i=1}^{N} \left[\lambda_i \|\eta_i\|^2 \left(A^{\mathrm{T}}P + PA - PBB^{\mathrm{T}}P\right)\right] \\ = & \frac{1}{2}\left(A^{\mathrm{T}}P + PA - PBB^{\mathrm{T}}P\right) \sum_{i=1}^{N} \lambda_i \|\eta_i\|^2, \end{aligned}$$

where the inequality results from (20). Considering the ARE in (19), the following can be obtained

$$V_1 + V_2 \leqslant -\frac{\beta}{2}\sum_{i=1}^{N} \lambda_i \|\eta_i\|^2 \leqslant -\frac{1}{2}\sum_{i=1}^{N} \lambda_i^2 \|\eta_i\|^2, \tag{28}$$

where the second inequality comes from (21).

Based on the well-known inequality

$$ab \leqslant \frac{r}{2}a^2 + \frac{1}{2r}b^2,$$

we can write V_3,

$$\begin{aligned} V_3 &= \sum_{i=1}^{N} c\lambda_i^2 \eta_i^{\mathrm{T}} PBB^{\mathrm{T}}P\varepsilon_i \\ &\leqslant \sum_{i=1}^{N} c\gamma\lambda_i^2 \left(\frac{r}{2}\|\eta_i\|^2 + \frac{1}{2r}\|\varepsilon_i\|^2\right), \end{aligned} \tag{29}$$

where $\gamma = \left\|PBB^{\mathrm{T}}P\right\|$ and r is a positive constant to be determined later.

Combining now (28) and (29) we have,

$$\begin{aligned} \dot{V} &\leqslant -\frac{1}{2}\sum_{i=1}^{N} \lambda_i^2 \|\eta_i\|^2 + \sum_{i=1}^{N} c\gamma\lambda_i^2 \left(\frac{r}{2}\|\eta_i\|^2 + \frac{1}{2r}\|\varepsilon_i\|^2\right) \\ &= \sum_{i=1}^{N} \frac{\lambda_i^2}{2}\left[\left(c\gamma r - 1\right)\|\eta_i\|^2 + \frac{c\gamma}{r}\|\varepsilon_i\|^2\right] \\ &\leqslant \frac{\beta^2}{2}\sum_{i=1}^{N} \left[\left(c\gamma r - 1\right)\|\eta_i\|^2 + \frac{c\gamma}{r}\|\varepsilon_i\|^2\right]. \end{aligned} \tag{30}$$

The second inequality comes from (21). By taking into account the orthogonal matrix U in (24) and (25), one has

$$\eta^{\mathrm{T}}\eta = \delta^{\mathrm{T}}\delta, \quad \varepsilon^{\mathrm{T}}\varepsilon = e^{\mathrm{T}}e.$$

Thus, (30) can be rewritten as

$$\begin{aligned} \dot{V} &\leqslant \frac{\beta^2}{2}\left[-\left(1 - c\gamma r\right)\|\delta\|^2 + \frac{c\gamma}{r}\|e\|^2\right] \\ &= \frac{\beta^2}{2}\left[-\left(1 - c\gamma r\right)\sum_{i=1}^{N}\|\delta_i\|^2 + \frac{c\gamma}{r}\sum_{i=1}^{N}\|e_i\|^2\right]. \end{aligned} \tag{31}$$

By adding and subtracting $\sigma_i\|\delta_i\|^2$, then (31) is equivalent to

$$\begin{aligned} \dot{V} \leqslant \frac{\beta^2\left(1 - c\gamma r\right)}{2}\sum_{i=1}^{N} \Bigg[& \frac{c\gamma}{r\left(1 - c\gamma r\right)}\|e_i\|^2 \\ & - \sigma_i\|\delta_i\|^2 - \left(1 - \sigma_i\right)\|\delta_i\|^2\Bigg], \end{aligned}$$

where, $\sigma_i \in (0, 1)$ is a design parameter $\forall i \in \mathcal{F}$.

In order to guarantee that $\dot{V} \leqslant 0$, the following conditions need to be satisfied $\forall i \in \mathcal{F}$

$$\|e_i\| \leqslant \pi_i \|\delta_i\|, \tag{32}$$

$$\pi_i = \sqrt{\sigma_i \frac{r}{c\gamma}\left(1 - c\gamma r\right)}, \tag{33}$$

where r is selected as $r < \frac{1}{c\gamma}$ to guarantee $1 - c\gamma r > 0$.

To this end, the event is triggered when (32) is violated, i.e.,

$$\|e_i\| \geqslant \pi_i \|\delta_i\|,$$

which guarantees that,

$$\dot{V} \leqslant \frac{\beta^2}{2}\left(1 - c\gamma r\right)\sum_{i=1}^{N}\left(\sigma_i - 1\right)\|\delta_i\|^2 \leqslant 0.$$

It should be noted that since $\delta_i = 0, \forall i \in \mathcal{F}$, this guarantees that each follower approaches the convex hull spanned by the leaders [13]. Therefore, $\dot{V} = 0$ if and only if the problem has a solution.

According to [30, Theorem 3.1], it follows that the solution set of $\dot{V} = 0$ is attractive and $\delta_i \to 0, \forall i \in \mathcal{F}$, as $t \to \infty$. □

Remark 2. The error $\delta_i(t)$ in (3) serves as a combinational measurement for containment control problem, which can be viewed as an extension of the combinational measurement in [31], [32] to the case of MASs with multiple leaders. □

Remark 3. The event-triggered containment control (12) depends on the coupling weight c and the feedback gain K. Theorem 1 gives the condition that the coupling gain c and the feedback gain K should satisfy. It can be seen that the feedback gain and coupling weight are decoupled and can be determined independently. It is shown in Section V that the feedback gain K will be obtained by RL. □

B. Feasibility of the Event-Triggered Mechanism

One critical issue in the event-triggered scheme design is to avoid Zeno behavior, i.e., the case where the minimal inter-event interval $\tau_{\min} = \min\limits_{j \in \mathbb{N}^+} \{t_{i+1} - t_i\}$ is zero and there is infinite number of events triggered over a finite time. In this section, the Zeno-free behavior of the proposed distributed event-triggered scheme in Theorem 1 is investigated as follows.

Theorem 2. *Consider the event-triggering condition designed by Theorem 1, then, $\forall i \in \mathcal{F}$, the inter-event intervals $\left\{t_{k+1}^i - t_k^i\right\}_{i=1}^{\infty}$ are strictly positive as $k \to \infty$.* □

Proof. We need to consider two cases separately $\forall i \in \mathcal{F}$.

i) The case when $\delta_i\left(t_k^i\right) \neq 0$.

We have proved in Theorem 1 that the event-triggering condition guarantees that $\delta \to 0$. The (32) is equivalent to

$$\left\|e_i(t)\right\|^2 \leqslant \frac{\pi_i^2\left(\left\|\delta_i(t)\right\|^2 + \left\|e_i(t)\right\|^2\right)}{1+\pi_i^2}. \quad (34)$$

Based on Young's inequality,

$$\frac{(a+b)^2}{2} \leqslant a^2 + b^2,$$

we have the following inequality,

$$\frac{\pi_i^2\left\|\delta_i(t)+e_i(t)\right\|^2}{2+2\pi_i^2} \leqslant \frac{\pi_i^2\left(\left\|\delta_i(t)\right\|^2 + \left\|e_i(t)\right\|^2\right)}{1+\pi_i^2}.$$

Therefore, the sufficient condition of (34) can be selected as

$$\begin{aligned}\left\|e_i(t)\right\|^2 &\leqslant \frac{\pi_i^2\left\|\delta_i(t)+e_i(t)\right\|^2}{2+2\pi_i^2} \\ &= \frac{\pi_i^2\left\|\hat{\delta}_i\left(t_k^i\right)\right\|^2}{2+2\pi_i^2} \triangleq s_k^i, \end{aligned} \quad (35)$$

To guarantee that $\delta \to 0$, an event is triggered, when (32) is violated. A sufficient condition for this is,

$$\left\|e_i(t)\right\| \leqslant \frac{\pi_i}{2+2\pi_i^2}\left\|\delta_i\left(t_k^i\right)\right\|.$$

The evolution of $e_i(t)$ over time in $\left[t_k^i, t_{k+1}^i\right)$ satisfies

$$\begin{aligned}\frac{d\left\|e_i(t)\right\|}{dt} &\leqslant \frac{\left\|e_i^{\mathrm{T}}(t)\right\|}{\left\|e_i(t)\right\|}\left\|\dot{e}_i(t)\right\| \\ &= \left\|\dot{\hat{\delta}}_i(t) - \dot{\delta}_i(t)\right\|, \forall t \in \left[t_k^i, t_{k+1}^i\right). \end{aligned} \quad (36)$$

The time derivative of $\delta_i(t)$ yields,

$$\begin{aligned}\dot{\delta}_i &= A\delta_i + Bu_i - c\sum_{j\in\mathcal{N}_i} a_{ij}BK\hat{\delta}_i(t) \\ &= A\delta_i - cBB^{\mathrm{T}}P\hat{\delta}_i(t) + cBB^{\mathrm{T}}P\sum_{j\in\mathcal{N}_i} a_{ij}\hat{\delta}_j\left(t_{k'(t)}^j\right). \end{aligned} \quad (37)$$

Based on (11), it is known that $\hat{\delta}_i(t)$ is constant for $\forall t \in \left[t_k^i, t_{k+1}^i\right)$. Taking (37) into (36) yields (38) and (39) (see next page), where

$$k'(t) = \arg\max\left\{t_k^j \middle| t_k^j \leqslant t, j \in \mathcal{N}_i\right\} \quad (40)$$

denotes the event-triggering time of follower j just before time t.

The summation in (38) stands for the previous update control value of all the neighbors of the follower i. The item $\frac{d\|e_i(t)\|}{dt}$ in (38) is the right-hand side derivative of $\left\|e_i(t)\right\|$ at the event-triggering time $t = t_k^i$.

Using the comparison lemma [33] in (38) yields,

$$\left\|e_i(t)\right\| \leqslant \frac{\omega_k^i}{\|A\|}\left(e^{\|A\|\left(t-t_k^i\right)} - 1\right). \quad (41)$$

where ω_k^i is defined in (39).

Combining now the inequalities (35) and (41) at the event-triggering instant t_{k+1}^i let us write,

$$\left\|e_i\left(t_{k+1}^i\right)\right\| = s_k^i \leqslant \frac{\omega_k^i}{\|A\|}\left(e^{\|A\|\left(t_{k+1}^i - t_k^i\right)} - 1\right),$$

where s_k^i is defined as in (35). Then,

$$\begin{aligned}\tau_k^i &= t_{k+1}^i - t_k^i \\ &\geqslant \frac{1}{\|A\|}\log\left(\frac{\|A\|\,s_k^i + \omega_k^i}{\omega_k^i}\right) \\ &> 0. \end{aligned}$$

ii) The case when $\delta_i\left(t_k^i\right) = 0$ as $k \to \infty$.

The sampling error (13) by using the absolute value inequality can be written as,

$$\left|\left\|\hat{\delta}_i(t)\right\| - \left\|\delta_i(t)\right\|\right| \leqslant \left\|e_i(t)\right\|. \quad (42)$$

Since the event-triggered condition designed by Theorem 1 guarantees (32) and combined with (42) gives,

$$\left|\left\|\hat{\delta}_i(t)\right\| - \left\|\delta_i(t)\right\|\right| \leqslant \pi_i\left\|\delta_i(t)\right\|,$$

which leads to

$$\frac{\left\|\delta_i\left(t_k^i\right)\right\|}{1+\pi_i} \leqslant \left\|\delta_i(t)\right\| \leqslant \frac{\left\|\delta_i\left(t_k^i\right)\right\|}{1-\pi_i}. \quad (43)$$

As a result of (43), we have,

$$\frac{\left\|\delta_i\left(t_k^i\right)\right\|}{\left\|\delta_i(t)\right\|} \geqslant 1 - \pi_i. \quad (44)$$

In the case when $\delta_i\left(t_k^i\right) = 0$ as $k \to \infty$, it follows from (43) that $\delta_i(t) = 0$. Then, the dynamics of the local containment error δ_i (37) satisfy

$$\begin{aligned}\dot{\delta}_i(t) &= A\delta_i(t) + cBB^{\mathrm{T}}P\sum_{j\in\mathcal{N}_i} a_{ij}\delta_j\left(t_{k'(t)}^j\right) \\ &\quad - cBB^{\mathrm{T}}P\delta_i(t) \\ &= 0, \end{aligned}$$

or,

$$A\delta_i = cBB^{\mathrm{T}}P\delta_i(t) - cBB^{\mathrm{T}}P\sum_{j\in\mathcal{N}_i} a_{ij}\delta_j\left(t_{k'(t)}^j\right). \quad (45)$$

Considering (39) and (45) yields

$$\begin{aligned}\omega_k^i &\leqslant \|A\|\left\|\delta_i\left(t_k^i\right)\right\| + \max_{t\in\left[t_k^i, t_{k+1}^i\right]}\left\|A\delta_i(t)\right\| \\ &= \|A\|\left\|\delta_i\left(t_k^i\right)\right\| + \left\|A\delta_i\left(t'\right)\right\|, \end{aligned}$$

where $t' \in \left[t_k^i, t_{k+1}^i\right]$. From (44), the following holds

$$\begin{aligned}\lim_{k\to\infty}\frac{s_k^i}{\omega_k^i} &\geqslant \lim_{k\to\infty}\frac{\frac{\pi_i}{\sqrt{2+2\pi_i^2}}\left\|\delta_i\left(t_k^i\right)\right\|}{\|A\|\left\|\delta_i\left(t_k^i\right)\right\| + \left\|A\delta_i\left(t'\right)\right\|} \\ &\geqslant \frac{(1-\pi_i)\,\pi_i}{\|A\|\,(2-\pi_i)\sqrt{2+2\pi_i^2}}. \end{aligned}$$

$$\frac{d\|e_i(t)\|}{dt} = \left\| A\left(\hat{\delta}_i(t) - e_i(t)\right) - cBB^{\mathrm{T}}P\hat{\delta}_i(t) + cBB^{\mathrm{T}}P\sum_{j\in\mathcal{N}_i} a_{ij}\hat{\delta}_j\left(t^j_{k'(t)}\right)\right\|$$

$$\leqslant \|Ae_i(t)\| + \left\| A\hat{\delta}_i(t) - cBB^{\mathrm{T}}P\hat{\delta}_i(t) + cBB^{\mathrm{T}}P\sum_{j\in\mathcal{N}_i} a_{ij}\hat{\delta}_j\left(t^j_{k'(t)}\right)\right\| \leqslant \|A\|\|e_i(t)\| + \omega^i_k, \tag{38}$$

$$\omega^i_k = \max_{t\in\left[t^i_k, t^i_{k+1}\right]} \left\| A\delta_i\left(t^i_k\right) - cBB^{\mathrm{T}}P\delta_i\left(t^i_k\right) + cBB^{\mathrm{T}}P\sum_{j\in\mathcal{N}_i} a_{ij}\hat{\delta}_j\left(t^j_{k'(t)}\right)\right\|, \tag{39}$$

Finally,

$$\begin{aligned}\lim_{k\to\infty}\tau^i_k &= \lim_{k\to\infty}\left(t^i_{k+1} - t^i_k\right)\\ &\geqslant \frac{1}{\|A\|}\log\left(\frac{\|A\|s^i_k}{\omega^i_k} + 1\right)\\ &\geqslant \frac{1}{\|A\|}\log\left(\frac{(1-\pi_i)\pi_i}{(2-\pi_i)\sqrt{2+2\pi_i^2}} + 1\right)\\ &> 0,\end{aligned}$$

which is the required result. □

V. Off-policy RL

Off-policy RL algorithms [34] are able to solve the ARE on-line in a model-free fashion. In this subsection, an off-policy RL algorithm is developed to solve (19) based on actor-critic structure [22].

Based on the discussions in Section IV, one can observe that the feedback gain $K = -B^TP$ depends on the ARE (19). In order to obtain the feedback gain $K = -B^TP$ in a model-free manner, the following problem is considered.

Problem 2. For each follower, design the optimal state feedback control $u_i = Kx_i$ such that the following performance function is minimized

$$\mathcal{V}^*_i(x_i) := \min_{u_i}\int_0^\infty \left(\beta\|x_i\|^2 + \|u_i\|^2\right)d\tau. \qquad \square$$

According to the optimal control theory [35], the optimal control that solves Problem 2 can be expressed as

$$u_i = Kx_i = -B^{\mathrm{T}}Px_i, \forall i\in\mathcal{F} \tag{46}$$

and the optimal performance is

$$\mathcal{V}^*_i(x_i) = x_i^{\mathrm{T}}Px_i, \forall i\in\mathcal{F}.$$

where P satisfies the ARE (19).

First, the optimal value function $\mathcal{V}^*_i(x_i)$ can be equivalently represented by a critic network as

$$\mathcal{V}^*_i(x_i) = \left(W^*_{c,i}\right)^{\mathrm{T}}\varphi_c(x_i),$$

with the weight vector $W^*_{c,i}\in\mathbb{R}^{\frac{n(n+1)}{2}}$ and quadratic polynomial basis vector

$$\varphi_c(x_i) = \begin{bmatrix} x_{i1}^2 & x_{i1}x_{i2} & \cdots & x_{in}^2\end{bmatrix}^{\mathrm{T}},$$

where $x_i := \begin{bmatrix} x_{i1} & \ldots & x_{in}\end{bmatrix}^{\mathrm{T}}$. By using approximators for $W^*_{c,i}$ as $W^\kappa_{c,i}$ in the κ-th iteration, the estimation of $\mathcal{V}^*_i(x_i)$ can be expressed as

$$\mathcal{V}^\kappa_i(x_i) = \left(W^\kappa_{c,i}\right)^{\mathrm{T}}\varphi_c(x_i), \tag{47}$$

and,

$$\nabla\mathcal{V}^\kappa_i(x_i) = \left[\nabla\varphi_c(x_i)\right]^{\mathrm{T}}W^\kappa_c.$$

Accordingly, the optimal control policy in (46) can be equivalently represented by an actor network in the form of

$$u^*_i(x_i) = \left(W^*_{a,i}\right)^{\mathrm{T}}\varphi_a(x_i),$$

where $\varphi_a(x_i) = x_i$ is the basis vector and $W^*_{a,i}\in\mathbb{R}^{n\times m}$ is the weights matrix of the actor network. In κ-th iteration, by denoting the estimation of $W^*_{a,i}$ as $W^\kappa_{a,i}$ one has

$$u^\kappa_i(x_i) = \left(W^\kappa_{a,i}\right)^{\mathrm{T}}\varphi_a(x_i). \tag{48}$$

Next, we need to rewrite the followers' dynamics (1) as,

$$\dot{x}_i = Ax_i + Bu^\kappa_i + B\left(m_i - u^\kappa_i\right), \ \forall i\in\mathcal{F}, \tag{49}$$

where m_i, $\forall i\in\mathcal{F}$ is the admissible policy[1] and $u^\kappa_i\in\mathbb{R}^{m\times 1}$ is the policy at κ-th iteration. Let the value function corresponding to the policy $u^\kappa_i(x)$ be written as

$$\mathcal{V}^\kappa_i(x_i) = x_i^{\mathrm{T}}P^\kappa_i x_i,$$

which after taking the time derivative along the system dynamics (49) becomes

$$\begin{aligned}\dot{\mathcal{V}}^\kappa_i &= \left\langle \frac{\partial\mathcal{V}^\kappa_i}{\partial x_i}, Ax_i + Bu^\kappa_i + B\left(m_i - u^\kappa_i\right)\right\rangle\\ &= x_i^{\mathrm{T}}\left(P^\kappa_i A + A^{\mathrm{T}}P^\kappa_i\right)x_i + 2x_i^{\mathrm{T}}P^\kappa_i Bu^\kappa_i\\ &\quad + 2x_i^{\mathrm{T}}P^\kappa_i B\left(m_i - u^\kappa_i\right),\end{aligned} \tag{50}$$

where $\langle\cdot,\cdot\rangle$ is the inner product operator. Integrating now both sides of (50) over an interval $[t, t+T]$ with $T\in\mathbb{R}^+$ yields the following off-policy integral RL Bellman equation,

$$\begin{aligned}&\mathcal{V}^\kappa_i(x_i(t+T)) - \mathcal{V}^\kappa_i(x_i(t))\\ &= \int_t^{t+T}\left(\beta\|x_i\|^2 + \|u_i\|^2\right)d\tau\\ &\quad + 2\int_t^{t+T} x_i^{\mathrm{T}}P^\kappa_i B\left(m_i - u^\kappa_i\right)d\tau.\end{aligned} \tag{51}$$

[1]For more details about admissible policy, see [35] for reference.

Denoting $v_i^{\kappa} := m_i - u_i^{\kappa}$, and considering the critic and actor networks (47) and (48) respectively, we can rewrite (51) in terms of the critic and actor weights $W_{c,i}^{\kappa}$ and $W_{a,i}^{\kappa+1}$ as

$$\begin{aligned}\varsigma_i^{\kappa} &= \left(W_{c,i}^{\kappa}\right)^{\mathrm{T}}\left[\varphi_c\left(x_i\left(t+T\right)\right) - \varphi_c\left(x_i\left(t\right)\right)\right] \\ &+ 2\sum_{j=1}^{m}\left(W_{a,i}^{\kappa+1,j}\right)^{\mathrm{T}}\int_t^{t+T}\varphi_a\left(x_i\right)v_i^{\kappa,j}d\tau \\ &- \int_t^{t+T}\left[\beta\|x_i\|^2 + \|\left(W_{a,i}^{\kappa}\right)^{\mathrm{T}}\varphi_a\left(x_i\right)\|^2\right]d\tau. \qquad (52)\end{aligned}$$

where

$$W_{a,i}^{\kappa+1} = \left[\begin{array}{ccccc} W_{a,i}^{\kappa+1,1} & \cdots & W_{a,i}^{\kappa+1,j} & \cdots & W_{a,i}^{\kappa+1,m}\end{array}\right],$$
$$v_i^{\kappa,j} = \left[\begin{array}{ccccc} v_i^{\kappa,1} & \cdots & v_i^{\kappa,j} & \cdots & v_i^{\kappa,m}\end{array}\right]^{\mathrm{T}}.$$

and ς_i^{κ} is a temporal difference error. By rearranging (52), the off-policy integral RL algorithm becomes a least squares problem of the form,

$$y_i^{\kappa}\left(t\right) = \left(W_i^{\kappa+1}\right)^{\mathrm{T}}\varphi_i\left(t\right) - \varsigma_i^{\kappa}\left(t\right), \qquad (53)$$

where,

$$y_i^{\kappa}\left(t\right) = \int_t^{t+T}\left[-\beta\|x_i\|^2 - \|\left(W_a^{\kappa}\right)^{\mathrm{T}}\varphi_a\left(x_i\right)\|^2\right]d\tau,$$
$$W_i^{\kappa+1} = \left[\begin{array}{cccc}\left(W_{c,i}^{\kappa}\right)^{\mathrm{T}} & \left(W_{a,i}^{\kappa+1,1}\right)^{\mathrm{T}} & \cdots & \left(W_{a,i}^{\kappa+1,m}\right)^{\mathrm{T}}\end{array}\right]^{\mathrm{T}}$$
$$\varphi_i\left(t\right) = \left[\begin{array}{c}\varphi_c\left(x_i\left(t+T\right)\right) - \varphi_c\left(x_i\left(t\right)\right) \\ 2\int_t^{t+T}\varphi_a\left(x_i\right)v_i^{\kappa,1}d\tau \\ \vdots \\ 2\int_t^{t+T}\varphi_a\left(x_i\right)v_i^{\kappa,m}d\tau\end{array}\right],$$

To this end, the off-policy RL algorithm, which solves the ARE (19) in a model-free manner, is given in Algorithm 2.

Algorithm 2: Off-policy RL Algorithm

Initialization: For follower i, set iteration index $\kappa = 0$, and start with an admissible control policy m_i;
procedure
  Least Squares Solution Step: Solve the least square problem (53) to obtain W_c^{κ} and $W_a^{\kappa+1}$ simultaneously;
  Stop if convergence is achieved, otherwise set $\kappa = \kappa + 1$ and go to Least Squares Solution Step;
  On convergence set $W_{a,i}^{*} = W_{a,i}^{\kappa}$.
end procedure

Remark 4. It can be seen that Algorithm 2 needs to start from an initial admissible policy. Due to page limitations, readers are referred to [23], [24], [34] for more details about how to obtain an initial admissible policy. □

VI. Simulation Study

Consider a MAS with three leaders and five followers with the communication graph illustrated in Figure 1. The system matrices of agent dynamics in (1) and (2) is selected as

$$A = \left[\begin{array}{cc}1 & -3 \\ 1 & -1\end{array}\right], B = \left[\begin{array}{c}0 \\ 1\end{array}\right].$$

Based on Theorem 1, the design parameters are selected as

$$\beta = 4.9985, c = 1.2713, \gamma = 25.3217, r = 0.0155.$$

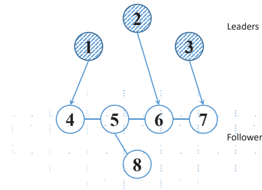

Fig. 1. Graph topology.

Then, the event-triggering condition (22) can be obtained with

$$\sigma_1 = 0.6, \sigma_2 = 0.5, \sigma_3 = 0.4, \sigma_4 = 0.3, \sigma_5 = 0.2$$
$$\pi_1 = 0.0120, \pi_2 = 0.0110, \pi_3 = 0.0098, \pi_4 = 0.0085,$$
$$\pi_5 = 0.0069.$$

The results of the event-triggered containment control are shown in Figure 2. One can observe that the state trajectories of all followers converge to the convex hull spanned by leaders. The event instants of all followers is shown in Figure 3, which gives the distributed and asynchronous feature of proposed event-triggered containment control protocols, because the event-triggering instant of each follower is independent of others.

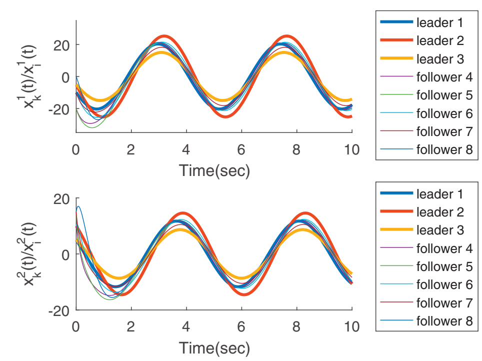

Fig. 2. State trajectories of agents.

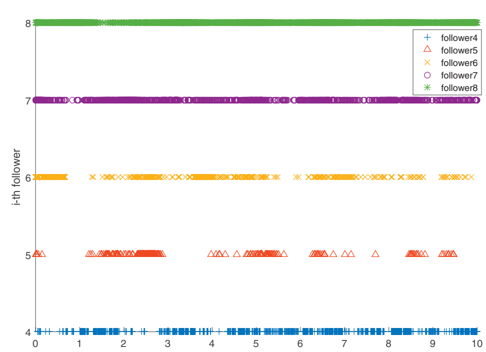

Fig. 3. The event instants of each follower.

VII. CONCLUSIONS

We have investigated the distributed event-triggered containment control design of MASs. The closed-loop stability of the equilibrium point for local containment error is analyzed. Moreover, Zeno-free behavior is guaranteed to be excluded. An off-policy RL algorithm is derived to obtain the feedback gain in the containment control without requiring the complete knowledge of the agents' dynamics. Simulation results shows the efficiency of the proposed approach.

Future efforts will focus on extending the results to more complicated cases, such as agents with nonlinear dynamics, time-delay in the communication between leaders and followers, directed communication graph, etc.

REFERENCES

[1] W. Ren, R. W. Beard, and E. M. Atkins, "Information consensus in multivehicle cooperative control," *IEEE Control Systems*, vol. 27, no. 2, pp. 71–82, 2007.

[2] Z. Kan, J. R. Klotz, E. L. Pasiliao, and W. E. Dixon, "Containment control for a social network with state-dependent connectivity," *Automatica*, vol. 56, pp. 86–92, 2015.

[3] J. A. Fax and R. M. Murray, "Information flow and cooperative control of vehicle formations," *IEEE transactions on automatic control*, vol. 49, no. 9, pp. 1465–1476, 2004.

[4] H. Su, X. Wang, and Z. Lin, "Flocking of multi-agents with a virtual leader," *IEEE Transactions on Automatic Control*, vol. 54, no. 2, pp. 293–307, Feb 2009.

[5] K.-K. Oh, M.-C. Park, and H.-S. Ahn, "A survey of multi-agent formation control," *Automatica*, vol. 53, pp. 424–440, 2015.

[6] K. G. Vamvoudakis, F. L. Lewis, and G. R. Hudas, "Multi-agent differential graphical games: Online adaptive learning solution for synchronization with optimality," *Automatica*, vol. 48, no. 8, pp. 1598–1611, 2012.

[7] J. Li, H. Modares, T. Chai, F. L. Lewis, and L. Xie, "Off-policy reinforcement learning for synchronization in multiagent graphical games," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 28, no. 10, pp. 2434–2445, Oct 2017.

[8] 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 Systems*, vol. 37, no. 1, pp. 33–52, Feb 2017.

[9] H. Modares, S. P. Nageshrao, G. A. D. Lopes, R. Babuška, and F. L. Lewis, "Optimal model-free output synchronization of heterogeneous systems using off-policy reinforcement learning," *Automatica*, vol. 71, pp. 334–341, 2016.

[10] B. Kiumarsi and F. L. Lewis, "Output synchronization of heterogeneous discrete-time systems: A model-free optimal approach," *Automatica*, vol. 84, no. Supplement C, pp. 86 – 94, 2017.

[11] Y. Cao, D. Stuart, W. Ren, and Z. Meng, "Distributed containment control for multiple autonomous vehicles with double-integrator dynamics: algorithms and experiments," *IEEE Transactions on Control Systems Technology*, vol. 19, no. 4, pp. 929–938, 2011.

[12] H. Liu, G. Xie, and L. Wang, "Necessary and sufficient conditions for containment control of networked multi-agent systems," *Automatica*, vol. 48, no. 7, pp. 1415–1422, 2012.

[13] H. Haghshenas, M. A. Badamchizadeh, and M. Baradarannia, "Containment control of heterogeneous linear multi-agent systems," *Automatica*, vol. 54, pp. 210–216, 2015.

[14] H. Liu, L. Cheng, M. Tan, and Z. Hou, "Containment control of general linear multi-agent systems with multiple dynamic leaders: A fast sliding mode based approach," *IEEE/CAA Journal of Automatica Sinica*, vol. 1, no. 2, pp. 134–140, 2014.

[15] G. Wen, Y. Zhao, Z. Duan, W. Yu, and G. Chen, "Containment of higher-order multi-leader multi-agent systems: a dynamic output approach," *IEEE Transactions on Automatic Control*, vol. 61, no. 4, pp. 1135–1140, 2016.

[16] P. Tabuada, "Event-triggered real-time scheduling of stabilizing control tasks," *IEEE Transactions on Automatic Control*, vol. 52, no. 9, pp. 1680–1685, 2007.

[17] X. Wang and M. D. Lemmon, "Event-triggering in distributed networked control systems," *IEEE Transactions on Automatic Control*, vol. 56, no. 3, pp. 586–601, 2011.

[18] D. V. Dimarogonas, E. Frazzoli, and K. H. Johansson, "Distributed event-triggered control for multi-agent systems," *IEEE Transactions on Automatic Control*, vol. 57, no. 5, pp. 1291–1297, 2012.

[19] E. Garcia, Y. Cao, and D. W. Casbeer, "Decentralized event-triggered consensus with general linear dynamics," *Automatica*, vol. 50, no. 10, pp. 2633–2640, 2014.

[20] T.-H. Cheng, Z. Kan, J. R. Klotz, J. M. Shea, and W. E. Dixon, "Decentralized event-triggered control of networked systems-part 2: Containment control," in *American Control Conference (ACC), 2015*. IEEE, 2015, pp. 5444–5448.

[21] W. Zhang, Y. Tang, Y. Liu, and J. Kurths, "Event-triggering containment control for a class of multi-agent networks with fixed and switching topologies," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 64, no. 3, pp. 619–629, 2017.

[22] K. G. Vamvoudakis and F. L. Lewis, "Online actor-critic algorithm to solve the continuous-time infinite horizon optimal control problem," *Automatica*, vol. 46, no. 5, pp. 878–888, 2010.

[23] H. Modares and F. L. Lewis, "Linear quadratic tracking control of partially-unknown continuous-time systems using reinforcement learning," *IEEE Transactions on Automatic control*, vol. 59, no. 11, pp. 3051–3056, 2014.

[24] H. Modares, F. L. Lewis, and Z. P. Jiang, "H_∞ tracking control of completely unknown continuous-time systems via off-policy reinforcement learning," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 26, no. 10, pp. 2550–2562, Oct 2015.

[25] R. Kamalapurkar, J. R. Klotz, P. Walters, and W. E. Dixon, "Model-based reinforcement learning in differential graphical games," *IEEE Transactions on Control of Network Systems*, 2017, in Press.

[26] F. L. Lewis, H. Zhang, K. Hengster-Movric, and A. Das, *Cooperative control of multi-agent systems: optimal and adaptive design approaches*. Springer Science & Business Media, 2013.

[27] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.

[28] H. Zhang, F. L. Lewis, and A. Das, "Optimal design for synchronization of cooperative systems: state feedback, observer and output feedback," *IEEE Transactions on Automatic Control*, vol. 56, no. 8, pp. 1948–1952, 2011.

[29] B. Paden and S. Sastry, "A calculus for computing filippov's differential inclusion with application to the variable structure control of robot manipulators," *IEEE Transactions on Circuits and Systems*, vol. 34, no. 1, pp. 73–82, Jan 1987.

[30] D. Shevitz and B. Paden, "Lyapunov stability theory of nonsmooth systems," *IEEE Transactions on Automatic Control*, vol. 39, no. 9, pp. 1910–1914, Sep 1994.

[31] 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, 2013.

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

[33] H. K. Khalil, *Noninear Systems*. Prentice-Hall, New Jersey, 1996.

[34] Y. Jiang and Z.-P. Jiang, "Computational adaptive optimal control for continuous-time linear systems with completely unknown dynamics," *Automatica*, vol. 48, no. 10, pp. 2699–2704, 2012.

[35] F. L. Lewis, D. Vrabie, and V. L. Syrmos, *Optimal control*. John Wiley & Sons, 2012.