Observer-based event-triggered optimal control for unknown nonlinear stochastic multi-agent systems with input constraints

Chen Liu
Hohai University
Lei Liu (liulei_hust@163.com)
Hohai University https://orcid.org/0000-0003-4529-5335
Zhaojing Wu
Yantai University
Jinde Cao
Southeast University School of Mathematics
Jianlong Qiu
Linyi University

Research Article

Keywords: Stochastic multi-agent systems, Adaptive critic designs, Event-triggered optimal control, Adaptive dynamic programming, Input constraints

Posted Date: June 1st, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1690634/v1

Noname manuscript No.
(will be inserted by the editor)

Observer-based event-triggered optimal control for unknown nonlinear stochastic multi-agent systems with input constraints

Chen Liu · Lei Liu · Zhaojing Wu · Jinde Cao · Jianlong Qiu

the date of receipt and acceptance should be inserted later

Abstract This paper addresses the observer-based event-triggered optimal control (ETOC) for nonlinear Itô-type stochastic multi-agent systems (SMASs) with unknown internal states and input constraints. To begin with, the event-triggered stochastic Hamilton-Jacobi-Bellman (HJB) equation with input constraints is presented for the first time. Next, the observer-based identifier network is utilized to recover the knowledge of unknown system dynamics. After that, the approximate event-triggered optimal controller is designed via adaptive critic designs (ACDs) whose weights are only updated at the triggering instants. It is worth mentioning that there is no published literature on the ETOC for nonlinear SMASs with unknown internal states and input constraints via the framework of ET-ACDs. This study is the first attempt to solve this problem. Moreover, it is also proved that the Zeno behavior does not exist in the closed-loop system. Finally, we present two examples to further verify the validity of the ETOC scheme.

C. Liu · L. Liu (✉)
College of Science, Hohai University, Nanjing 210098, China
e-mail: liulei_hust@163.com

Z. Wu
School of Mathematics and Informational Science, Yantai University, Yantai 264005, China

J. Cao
School of Mathematics, Southeast University, Nanjing 210096, China, and also with Yonsei Frontier Lab, Yonsei University, Seoul 03722, South Korea

J. Qiu
School of Automation and Electrical Engineering, Linyi University, Linyi 276005, China

Keywords Stochastic multi-agent systems · Adaptive critic designs · Event-triggered optimal control · Adaptive dynamic programming · Input constraints

1 Introduction

Recently, the distributed coordination control of nonlinear multi-agent systems (MASs) has been widely used in various fields, such as the manipulator works in transportation equipment; the flocking phenomenon of biological systems; the formation control of UAVs; the electronic circuit system. For more details, see [1–4] and references therein. Actually, these fields are inevitably affected by stochastic factors. Therefore, it is essential to study the distributed coordination control of SMASs, which has also aroused the interest of many researchers [5–9].

As a typical topic in distributed coordination control of SMASs, the mean-square leader-following consensus implies that all followers reach the leader in mean-square sense by appropriate control strategies [7]. Ren et al. [8] and Wei et al. [9] studied the mean-square leader-following consensus problem for SMASs under the undirected and directed topology, respectively. Similar to [5–9], the existing research results are mostly focused on the consensus. However, optimality, as another important topic of SMASs, has not been paid enough attention, which refers to minimizing the consumption of the performance index on the basis of achieving consensus. As we all know, the dynamic programming is a method to study the optimal control problem by defining a Hamilton-Jacobi-Bellman (HJB) equation. While, in practice, the use of the dynamic programming algorithm to solve the analytical solution of the HJB equation has great limitations. To overcome these shortcomings, Werbos et al. [10] designed the adaptive dynamic programming (ADP) algorithm by utilizing the idea of function approximation for the first time and obtained the numerical solution of the HJB equation. Since both have almost the same principles, ADP is also called ACDs. In recent years, many novel variants of the ADP algorithm have been proposed. To work out the optimal coordination control problem of nonlinear MASs, a fuzzy ADP algorithm was designed by combining generalized fuzzy hyperbolic model with adaptive dynamic programming in [11]. Later, a novel online learning consensus control strategy for MASs was developed via a goal representation heuristic dynamic programming (GrHDP) techniques in [12]. For more details on the ADP algorithm, see [13–16] and references therein.

It is well known that time-triggered control strategies often lead to a large number of unnecessary communication transmissions, which seriously reduces communication efficiency. To overcome this shortcoming, the event-triggered control (ETC) mechanism was proposed, which can effectively reduce controller updates and redundant communication between plants and actuators by designing an aperiodic method to transmit the system signals [17]. Therefore, many studies combined the ADP with the ETC mechanism [18–21]. A novel optimal adaptive ETC algorithm for continuous-time (CT) systems was proposed via the actor-critic algorithm in [18]. Later, to simplify the architecture of the algorithm, Zhao et al. [19] and Sun et al. [20] utilized the critic-only network to design the ETOC for nonlinear MASs without and with external disturbances, respectively. The ETOC problem of nonlinear stochastic systems was sudied for the first time by utilizing the ADP in [21], in which event-triggered conditions strictly rely on the known system dynamics, and there is no proof that the Zeno behavior has been excluded.

Due to the physical characteristics of actuators in engineering applications, actuator saturation (i.e., input constraints) needs to be considered [22–25]. To tackle this issue, Dong et al. [22] combined the ETC mechanism with the actor-critic algorithm, designing the event-triggered optimal controller of the nonlinear system with input constraints. After that, by using the critic-only network, Wang et al. [23] and Yang et al. [24] focused on establishing the robust ETC strategy for constrained-input nonlinear systems with matched and mismatched perturbations, respectively. Recently, Shi et al. [25] applied the ET-ACDs to investigate the optimal control problem of the CT nonlinear MASs with input saturation.

Note that all the aforementioned results rely on the full knowledge of system dynamics. Actually, it is difficult to acquire knowledge of the system dynamics in practical applications completely. To handle with the unknown nonlinear dynamics, partially models or model-free based event-triggered adaptive dynamic programming (ET-ADP) methods have been extensively studied [26–30]. Zhu et al. [26] proposed a novel ADP algorithm by combining the identifier network with the actor-critic network to design the ETOC of partially unknown constrained-input nonlinear systems. Thereafter, to simplify the architecture of the algorithm, Zhang et al. [27] and Huo et al. [28] applied the identifier-critic framework to study the robust event-triggered control strategy for unknown constrained-input nonlinear systems with matched and mismatched perturbations, respectively. Recently, for unknown MASs, Zhang et al. [29] addressed the optimal control problem with input constraints by utilizing the ET-ADP technique. More recently, Ding et al. [30] studied the optimal control problem of CT nonlinear MASs with input constraints by designing a new neural-network-based observer and exploiting an actor-critic network.

Inspired by the aforementioned discussions, it should be pointed out that the corresponding ETOC problem of SMASs via the framework of ET-ACDs has not been fully studied, let alone considering the case model-free and input constraints at the same time. It motivates our research interest and the main contributions of this paper are:

(1) The ETOC problem for nonlinear SMASs with unknown internal states and input constraints is investigated for the first time, and the ETOC strategy induced by the event-triggered stochastic HJB equation of SMASs with input constraints is presented via the Itô formula and Bellman's optimality principle.

(2) A sufficient criterion on optimal mean-square leader-following consensus of SMASs with input constraints via the ETOC strategy is derived and it is also proved that the closed-loop system can exclude the Zeno behavior.

(3) The observer-based identifier network is utilized to reconstruct the unknown system dynamics and the approximate event-triggered optimal controller of SMASs with input constraints is designed by utilizing the framework of ET-ACDs. It is worth emphasizing that all error signals in the above two networks are semi-globally uniformly ultimately bounded (SGUUB) in mean-square sense.

The rest of the paper is planned as follows. The problem formulation is described in Sect. 2. Sect. 3 gives the ETOC strategy with input constraints. Sect. 4 presents the system identification and the design of approximate event-triggered optimal controller of SMASs. Sect. 5 and 6 are the simulation and the conclusion, respectively.

Notations: The set of all nonnegative real numbers, the n-dimensional Euclidean space, the $m \times n$ real matrices, and the n-dimensional identity matrix are denoted by $\mathbb{R}^+$, $\mathbb{R}^n$, $\mathbb{R}^{m\times n}$, and I_n, respectively. $\underline{\mathbf{1}}_n = [1, \ldots, 1]^T \in \mathbb{R}^n$. $\|\cdot\|$ represents the two-norm for vectors or corresponding indeced two-norm for matrices. $M > 0 (M < 0)$ denotes M is symmetric positive (negative) definite. $\mathbb{C}^{2,1}$ represents a class of functions $V(x,t)$ on $\mathbb{R}^n \times \mathbb{R}^+ \to \mathbb{R}^+$ with twice continuously differentiable in x and once in t. The minimum (maximum) eigenvalue of matrix M is denoted by $\lambda_{\min}(M)$ ($\lambda_{\max}(M)$).

2 Preliminaries

2.1 Algebraic graph theory

The directed graph can be described as $\mathcal{G} = (\Pi, \xi, \mathcal{A})$, where $\Pi = \{\pi_0, \pi_1, \pi_2, \cdots, \pi_N\}$ represents the node-set, $\xi = \{\tilde{e}_{ij} = (\pi_i, \pi_j)\} \subseteq \Pi \times \Pi$ denotes the edge-set, and $\mathcal{A} = [a_{ij}] \in \mathbb{R}^{N\times N}$ represents the weighted adjacency matrix with $a_{ij} \geq 0$. The element $a_{ij} > 0$ only if $\tilde{e}_{ji} \in \xi$, which means π_i can receive the information from π_j; otherwise $a_{ij} = 0$. $N_i = \{\pi_j | \tilde{e}_{ji} \in \xi\}$ denotes the neighbor set of node π_i. Define the pinning matrix of leader as $\mathcal{B} = \mathrm{diag}\{b_i, i = 1, 2, \cdots, N\}$ where b_i denotes the pinning gain from π_0 to π_i. Let the in-degree matrix of $\mathcal{G}$ be $\mathcal{D} = \mathrm{diag}\{d_i, i = 1, 2\cdots, N\}$ with $d_i = \sum_{j\in\mathcal{N}_i} a_{ij}$. Then, the corresponding digraph Laplacian matrix is defined as $\mathcal{L} = \mathcal{D} - \mathcal{A} = [l_{ij}] \in \mathbb{R}^{N\times N}$ with $l_{ii} = \sum_{j\in\mathcal{N}_i} a_{ij}$ and $l_{ij} = -a_{ij} (i \neq j)$. Let $\mathcal{K} = \mathcal{L} + \mathcal{B}$ for convenience.

2.2 Problem formulation

For any $(x_{is}, s) \in \mathbb{R}^n \times [0, +\infty)$, consider following nonlinear SMASs during $t \in [s, +\infty)$:

$$\begin{cases} dx_i(t) = f(x_i(t), t)dt + u_i(t)dt + g(x_i(t), t)dw(t) \\ y_i(t) = Cx_i(t) \end{cases} \tag{1}$$

where $x_i(t) \in \mathbb{R}^n$ and $u_i(t) \in \mathbb{R}^n$ are the system state and control input of agent i, respectively; the initial data is defined as $x_i(s) = x_{is}$; $y_i(t) \in \mathbb{R}^p$ is the output vector of agent i; $f : \mathbb{R}^n \times \mathbb{R}^+ \to \mathbb{R}^n$ and $g : \mathbb{R}^n \times \mathbb{R}^+ \to \mathbb{R}^{n\times k}$ are both unknown smooth functions with $f(0,t) = g(0,t) = 0$; $C \in \mathbb{R}^{p\times n}$ is the known output matrix. Let $w(t) = (w_1(t), \cdots, w_k(t))^T$ be a k-dimensional normal Brownian motion defined on a complete probability space $(\Omega, \mathscr{F}, \{\mathscr{F}_t\}_{t\geq s}, \mathbb{P})$. $\mathbb{E}[\cdot]$ represents the mathematical expectation of $\mathbb{P}$.

The dynamics of the leader is defined as

$$dx_0(t) = f(x_0(t), t)dt + g(x_0(t), t)dw(t). \tag{2}$$

Definition 1 [31] Consider a stochastic nonlinear system

$$dx(t) = f(x(t), t)dt + g(x(t), t)dw(t) \tag{3}$$

for any $V(x,t) \in \mathbb{C}^{2,1}$, associated with system (3), the differential operator $\mathscr{L}$ is defined as:

$$\mathscr{L}V(x,t) = V_t(x,t) + V_x(x,t)f + \frac{1}{2}\mathrm{Tr}(g^T V_{xx} g) \tag{4}$$

where $V_t(x,t) = \frac{\partial V}{\partial t}$, $V_x(x,t) = \frac{\partial V}{\partial x}$, $V_{xx}(x,t) = \frac{\partial^2 V}{\partial x^2}$, and $\mathrm{Tr}(\cdot)$ is the matrix trace.

Definition 2 [32] Considering the system (3), its trajectory $x(t)$ is said to be SGUUB in pth moment, if for any initial data (x_s, s), one has $\mathbb{E}[\|x(t)\|^p] < \varepsilon$ for all $t > s + T(\varepsilon, x_s)$ where ε and $T(\varepsilon, x_s)$ are positive constants. When $p = 2$, it is also called SGUUB in mean-square sense.

Lemma 1 [33] *Let $V : \mathbb{R}^n \times \mathbb{R}^+ \to \mathbb{R}^+$ be a continuous function and $V \in \mathbb{C}^{2,1}$. If there exist positive constants a, b and two class k_∞ functions $\alpha_1(\cdot), \alpha_2(\cdot)$ such that*

$$\alpha_1(\|x\|) \leq V(x,t) \leq \alpha_2(\|x\|) \tag{5}$$

$$\mathscr{L}V(x,t) \leq -aV(x,t) + b. \tag{6}$$

Then, it can be derived that

$$\mathbb{E}[V(x,t)] \leq V(x(0), 0)e^{-at} + \frac{b}{a} \tag{7}$$

for each $x(0) \in \mathbb{R}^n$, $\forall t \geq 0$. It indicates that the solution $x(t)$ is SGUUB in mean-square sense.

Assumption 1 [34] For the directed graph $\mathcal{G}$ of SMASs (1), the leader π_0 has a directed path (the root) to any other follower $\pi_i \in \Pi$ $(i = 1, \ldots, N)$.

Assumption 2 [35] Assume that it exists a positive constant L_f such that $\|f(x,t) - f(y,t)\| \leq L_f\|x - y\|$ for all $x, y \in \mathbb{R}^n$. Moreover, $g^T g$ is bounded, i.e., $\|g^T g\| \leq g_M$ where g_M is a positive constant.

3 The ETOC of SMASs with input constraints

3.1 The stochastic HJB equation with input constraints

Let $e_i(t) = x_i(t) - x_0(t)$. For any $t \in [s, +\infty)$, the consensus error system of agent i can be written as:

$$de_i(t) = \tilde{f}(e_i(t), t)dt + u_i(t)dt + \tilde{g}(e_i(t), t)dw(t). \tag{8}$$

Let $x(t)=[x_1^T(t),\cdots,x_N^T(t)]^T\in\mathbb{R}^{Nn}$, $e(t)=[e_1^T(t),\cdots,e_N^T(t)]^T\in\mathbb{R}^{Nn}$, $F(e(t),t)=[\tilde{f}^T(e_1(t),t),\cdots,\tilde{f}^T(e_N(t),t)]^T\in\mathbb{R}^{Nn}$ where $\tilde{f}(e_i(t),t)=f(x_i(t),t)-f(x_0(t),t)$, $u(t)=[u_1^T(t),\cdots,u_N^T(t)]^T\in\mathbb{R}^{Nn}$, $G(e(t),t)=[\tilde{g}^T(e_1(t),t),\cdots,\tilde{g}^T(e_N(t),t)]^T\in\mathbb{R}^{Nn\times k}$ where $\tilde{g}(e_i(t),t)=g(x_i(t),t)-g(x_0(t),t)$. Then, for any $t\in[s,+\infty)$, the global consensus error system becomes

$$de(t)=F(e(t),t)dt+u(t)dt+G(e(t),t)dw(t). \tag{9}$$

For any initial data $(z_i,s)\in\mathbb{R}^n\times[0,+\infty)$, the local neighbor error system of agent i during $t\in[s,+\infty)$ is depicted by

$$\begin{cases} d\zeta_i(t)=\sum\limits_{j\in\mathcal{N}_i}a_{ij}(dx_i(t)-dx_j(t))+b_i(dx_i(t)-dx_0(t)) \\ \zeta_i(s)=z_i. \end{cases} \tag{10}$$

Then, the global neighbor error is defined as

$$\begin{aligned} d\zeta(t)&=(\mathcal{L}\otimes I_n)dx(t)+(\mathcal{B}\otimes I_n)(dx(t)-d\underline{x}_0(t)) \\ &=(\mathcal{L}\otimes I_n)(dx(t)-d\underline{x}_0(t))+(\mathcal{B}\otimes I_n)de(t) \\ &=(\mathcal{K}\otimes I_n)de(t) \end{aligned} \tag{11}$$

where $\underline{x}_0(t)=\left[x_0^T(t),\cdots,x_0^T(t)\right]^T\in\mathbb{R}^{Nn}$ and $\otimes$ denotes the Kronecker product.

Remark 1 As the Laplacian matrix $\mathcal{L}$ has at least a zero eigenvalue and the corresponding eigenvector is $\underline{\mathbf{1}}_N$, i.e., $\mathcal{L}\underline{\mathbf{1}}_N=0$, (10) can be derived to (11). Additionally, $e(t)\to 0$ is equivalent to $\zeta(t)\to 0$ if $\mathcal{K}$ is nonsingular.

Definition 3 [36] If the control input u_i guarantees the performance index finite on the basis of making the SMASs (1) stable, then, u_i is called an admissible control. We write as $u_i\in\mathcal{U}_i[s,+\infty)$ where $\mathcal{U}_i[s,+\infty)=\{u_i(\cdot)$ is measurable in $[s,+\infty)$, and $\{\mathscr{F}_t\}_{t\geq s}$-adapted$\}$ is the admissible control set.

For any $\hat{s}\in[s,+\infty)$ and $\{\mathscr{F}_{\hat{s}}\}_{\hat{s}\geq s}$-measurable random variable $\hat{z}_i$, the local non-quadratic performance index of agent i is define as

$$J_i(\hat{z}_i,\hat{s};u_i,u_{(i)})=\mathbb{E}_{\hat{s}}\left[\int_{\hat{s}}^{+\infty}\left(\zeta_i^TQ_{ii}\zeta_i+M_i(u_i)\right)dt\right] \tag{12}$$

where $\mathbb{E}_{\hat{s}}[\cdot]$ is the conditional expectation of $\{\mathscr{F}_{\hat{s}}\}_{\hat{s}\geq s}$, $\|u_i\|\leq\bar{u}_i$ with $\bar{u}_i$ being a positive constant, $u_{(i)}$ denotes a set containing all neighbor control vectors of agent i, $M_i(u_i)=2\bar{u}_i\int_0^{u_i}\tanh^{-T}(v_i/\bar{u}_i)R_{ii}dv_i$ is positive definite with $Q_{ii}>0$, $R_{ii}>0$.

Definition 4 (Optimal Control [36]) For given initial data (z_i,s), if there is an admissible control $u_i^*\in\mathcal{U}_i[s,+\infty)$ minimizes (12), we say that u_i^* is an optimal control. The associated state trajectory is an optimal state at (z_i,s).

For agent i, let local optimal value function $V_i^*(z_i,s)$ be

$$V_i^*(z_i,s)=\min_{u_i\in\mathcal{U}_i[s,+\infty)}J_i(z_i,s;u_i,u_{(i)}). \tag{13}$$

Based on stochastic optimal control theory [37], the corresponding local stochastic HJB equation becomes

$$\begin{aligned} -V_{i,t}^*=&\min_{u_i\in\mathcal{U}_i[s,+\infty)}(\zeta_i^TQ_{ii}\zeta_i+M_i(u_i) \\ &+V_{\zeta_i}^{*T}(\mathcal{K}_i\otimes I_n)(F(e(t),t)+u(t)) \\ &+\frac{1}{2}\mathrm{Tr}\left(G^T(\mathcal{K}_i^T\otimes I_n)\frac{\partial^2V_i^*}{\partial\zeta_i^2}(\mathcal{K}_i\otimes I_n)G\right) \\ =&\min_{u_i\in\mathcal{U}_i[s,+\infty)}(\zeta_i^TQ_{ii}\zeta_i+2\bar{u}_i\int_0^{u_i}\tanh^{-T}(v_i/\bar{u}_i)R_{ii}dv_i \\ &+V_{\zeta_i}^{*T}((l_{ii}+b_{ii})\otimes I_n)(\tilde{f}_i+u_i(t)) \\ &+V_{\zeta_i}^{*T}\sum_{j\in\mathcal{N}_i}((l_{ij}+b_{ij})\otimes I_n)(\tilde{f}_j+u_j(t)) \\ &+\frac{1}{2}\mathrm{Tr}\left(G^T(\mathcal{K}_i^T\otimes I_n)\frac{\partial^2V_i^*}{\partial\zeta_i^2}(\mathcal{K}_i\otimes I_n)G\right) \end{aligned} \tag{14}$$

where $V_{i,t}^*=\frac{\partial V_i^*(\zeta_i,t)}{\partial t}$, $V_{\zeta_i}^*=\frac{\partial V_i^*(\zeta_i,t)}{\partial\zeta_i}$, $\tilde{f}_i=\tilde{f}(e_i(t),t)$, $G^T(\mathcal{K}_i^T\otimes I_n)\frac{\partial^2V_i^*}{\partial\zeta_i^2}(\mathcal{K}_i\otimes I_n)G\in\mathbb{R}^{k\times k}$, $\mathcal{K}_i=\mathcal{L}_i+\mathcal{B}_i$, with $\mathcal{L}_i$ and $\mathcal{B}_i$ being the ith row vectors of matrix $\mathcal{L}$ and $\mathcal{B}$, respectively. Simultaneously, the corresponding local Hamilton function can be denoted by

$$\begin{aligned} &H_i(\zeta_i,V_{\zeta_i},u_i,u_{(i)},t) \\ &=\zeta_i^TQ_{ii}\zeta_i+2\bar{u}_i\int_0^{u_i}\tanh^{-T}(v_i/\bar{u}_i)R_{ii}dv_i \\ &\quad+V_{\zeta_i}^T(\mathcal{K}_i\otimes I_n)\left(F(e(t),t)+u(t)\right) \\ &\quad+\frac{1}{2}\mathrm{Tr}\left(G^T(\mathcal{K}_i^T\otimes I_n)\frac{\partial^2V_i}{\partial\zeta_i^2}(\mathcal{K}_i\otimes I_n)G\right). \end{aligned} \tag{15}$$

Then, (14) can be rwritten as

$$-V_{i,t}^*=\min_{u_i\in\mathcal{U}_i[s,+\infty)}H_i\left(\zeta_i,V_{\zeta_i}^*,u_i,u_{(i)},t\right). \tag{16}$$

From (16) and Bellman's optimality principle, one gets

$$u_i^*(t)=-\bar{u}_i\tanh(D_i^*) \tag{17}$$

where $D_i^*=\frac{1}{2\bar{u}_i}R_{ii}^{-1}((l_{ii}+b_{ii})\otimes I_n)V_{\zeta_i}^*$. By (17), we have

$$\begin{aligned} M_i(u_i^*)=&\bar{u}_iV_{\zeta_i}^{*T}((l_{ii}+b_{ii})\otimes I_n)\tanh(D_i^*) \\ &+\bar{u}_i^2\bar{R}_{ii}\ln(\underline{\mathbf{1}}_n-\tanh^2(D_i^*)) \end{aligned} \tag{18}$$

where $\bar{R}_{ii}=[r_1,\cdots,r_n]\in\mathbb{R}^{1\times n}$. Substituting (17) and (18) into (16), the CT stochastic HJB equation of agent i becomes

$$\begin{aligned} &V_{i,t}^*+H_i(\zeta_i,V_{\zeta_i}^*,u_i^*,u_{(i)}^*,t) \\ &\quad=V_{i,t}^*+\zeta_i^TQ_{ii}\zeta_i+V_{\zeta_i}^{*T}((l_{ii}+b_{ii})\otimes I_n)\tilde{f}_i \end{aligned}$$

$$
\begin{aligned}
&+\bar{u}_i^2\bar{R}_{ii}\ln(\underline{\mathbf{1}}_n-\tanh^2(D_i^*))\\
&+V_{\zeta_i}^{*T}\sum_{j\in\mathcal{N}_i}((l_{ij}+b_{ij})\otimes I_n)(\tilde{f}_j-\bar{u}_j\tanh(D_j^*))\\
&+\frac{1}{2}\mathrm{Tr}\left(G^T(\mathcal{K}_i^T\otimes I_n)\frac{\partial^2 V_i^*}{\partial\zeta_i^2}(\mathcal{K}_i\otimes I_n)G\right)\\
=&\ 0. \qquad (19)
\end{aligned}
$$

To reduce controller updates and redundant communication between the plants and the actuators, the ETC mechanism is adopted in this paper. Define an event-triggered sequence of agent i as $\{t_k^i\}$. For any $t\in[t_k^i,t_{k+1}^i)$, the measurement error is $r_i(t)=\zeta_i(t_k^i)-\zeta_i(t)$ and $r(t)=[r_1^T(t),\cdots,r_N^T(t)]^T\in\mathbb{R}^{Nn}$. Meanwhile, the ETOC can be obtained during $t\in[t_k^i,t_{k+1}^i)$

$$
u_i^*(t_k^i)=-\bar{u}_i\tanh\big(\frac{1}{2\bar{u}_i}R_{ii}^{-1}((l_{ii}+b_{ii})\otimes I_n)V_{\zeta_i}^*(t_k^i)\big). \qquad (20)
$$

Substituting (20) into (16), the event-triggered stochastic HJB equation becomes

$$
\begin{aligned}
&V_{i,t}^*+H_i(\zeta_i,V_{\zeta_i}^*,u_i^*(t_k^i),u_{(i)}^*,t)\\
&\quad=V_{i,t}^*+\zeta_i^TQ_{ii}\zeta_i+M_i(u_i^*(t_k^i))\\
&\quad\quad+V_{\zeta_i}^{*T}((l_{ii}+b_{ii})\otimes I_n)(\tilde{f}_i-\bar{u}_i\tanh(D_i^*(t_k^i)))\\
&\quad\quad+V_{\zeta_i}^{*T}\sum_{j\in\mathcal{N}_i}((l_{ij}+b_{ij})\otimes I_n)(\tilde{f}_j-\bar{u}_j\tanh(D_j^*))\\
&\quad\quad+\frac{1}{2}\mathrm{Tr}\left(G^T(\mathcal{K}_i^T\otimes I_n)\frac{\partial^2 V_i^*}{\partial\zeta_i^2}(\mathcal{K}_i\otimes I_n)G\right)\\
&\quad=\varepsilon_{ui} \qquad (21)
\end{aligned}
$$

where ε_{ui} is the transformation error generated from (17) to (20). By (19) and (21), we have $\varepsilon_{ui}=M_i(u_i^*(t_k^i))-M_i(u_i^*)+2\bar{u}_iD_i^{*T}R_{ii}(u_i^*(t_k^i)-u_i^*)$. Furthermore, it can be observed that the event-triggered stochastic HJB equation (21) is equal to zero only at the triggering instants.

Remark 2 For unconstrained optimal control problems, the performance index is usually quadratic, i.e., $M_i(u_i)=u_i^TR_{ii}u_i$ where $R_{ii}>0$. For constrained-input optimal control problems, the performance index is nonquadratic, i.e., $M_i(u_i)=2\bar{u}_i\int_0^{u_i}\tanh^{-T}(v_i/\bar{u}_i)R_{ii}dv_i>0$, which ensures that the control input is within an interval. Compared with the unconstrained optimal control problem, the constrained-input optimal control problem has brought more difficulties to theoretical analysis and numerical simulation.

Remark 3 The HJB equations of deterministic systems have been fully studied in the past several years. Additionally, the coupled Hamilton-Jacobi-Isaacs (HJI) equations of zero-sum differential graphical games have also been studied by many scholars [38–40]. However, as far as we know, optimality, as another important topic of SMASs, has not been adequately studied. In this paper, by using the Itô formula and Bellman's optimality principle, the stochastic HJB equation (21) of SMASs with input constraints under ETC mechanism is presented for the first time, in which the second-order differential $\frac{\partial^2 V_i^*}{\partial\zeta_i^2}$ makes the design of the controller much more difficult than that of the deterministic case. Therefore, the study of ETOC for unknown nonlinear SMASs with input constraints is a meaningful and challenging work.

3.2 Event-triggered optimal control strategy

Assumption 3 [26] Assume that there exists a positive constant L_{D_i} such that $\|D_i^*(x_1)-D_i^*(x_2)\|\leq L_{D_i}\|x_1-x_2\|$ for all $x_1,x_2\in\mathbb{R}^n$.

Assumption 4 For optimal value funtion $V_i^*(\zeta_i,t):\mathbb{R}^n\times\mathbb{R}^+\to\mathbb{R}^+$, suppose that there exist two positive scalars c_1,c_2 such that $c_1\mathbb{E}[\|\zeta_i\|^2]\leq\mathbb{E}[V_i^*(\zeta_i,t)]\leq c_2\mathbb{E}[\|\zeta_i\|^2]$.

Theorem 1 *Let Assumptions 1-4 valid and $V_i^*(\zeta_i,t)$ be the solution of the stochastic HJB equation* (19). *Then, the SMASs* (1) *will achieve the optimal mean-square leader-following consensus under ETOC* (20) *and following triggering condition*

$$
\mathbb{E}[\|r_i(t)\|^2]\leq\frac{(1-\chi^2)\lambda_{\min}(Q_{ii})}{\bar{u}_i^2L_{D_i}^2\|R_{ii}\|}\mathbb{E}[\|\zeta_i\|^2]+\frac{e^{-\epsilon t}}{\bar{u}_i^2L_{D_i}^2\|R_{ii}\|} \qquad (22)
$$

for $t\in[t_k^i,t_{k+1}^i)$, where $0<\chi<1$ and $\epsilon>0$ are event-triggered parameters. Moreover, it is also proved that the closed-loop system can exclude the Zeno behavior.

Proof Step 1. The optimal mean-square leader-following consensus of SMASs (1) under the ETC mechanism can be achieved. As the optimal value function $V_i^*(\zeta_i,t)$ is positive definite, it can be selected as the Lyapunov function. By Definition 1, one has

$$
dV_i^*(\zeta_i,t)=\mathscr{L}V^*(\zeta_i,t)dt+V_{\zeta_i}^{*T}(\mathcal{K}_i\otimes I_n)Gdw(t) \qquad (23)
$$

where

$$
\begin{aligned}
\mathscr{L}V_i^*(\zeta_i,t)=&\ V_{i,t}^*+V_{\zeta_i}^{*T}((l_{ii}+b_{ii})\otimes I_n)(\tilde{f}_i+u_i^*(t_k^i))\\
&+V_{\zeta_i}^{*T}\sum_{j\in\mathcal{N}_i}((l_{ij}+b_{ij})\otimes I_n)(\tilde{f}_j+u_j^*(t))\\
&+\frac{1}{2}\mathrm{Tr}\left(G^T(\mathcal{K}_i^T\otimes I_n)\frac{\partial^2 V_i^*}{\partial\zeta_i^2}(\mathcal{K}_i\otimes I_n)G\right). \qquad (24)
\end{aligned}
$$

According to CT stochastic HJB equation (19), we have

$V_{\zeta_i}^{*T}((l_{ii}+b_{ii})\otimes I_n)\tilde{f}_i$

$$
\begin{aligned}
&= -V_{i,t}^* - \zeta_i^T Q_{ii}\zeta_i - \bar{u}_i^2 \bar{R}_{ii} \ln\left(\underline{\mathbf{1}}_n - \tanh^2(D_i^*)\right) \\
&\quad - V_{\zeta_i}^{*T} \sum_{j\in\mathcal{N}_i} ((l_{ij}+b_{ij})\otimes I_n)(\tilde{f}_j - \bar{u}_j \tanh(D_j^*)) \\
&\quad - \frac{1}{2}\mathrm{Tr}\left(G^T(\mathcal{K}_i^T\otimes I_n)\frac{\partial^2 V_i^*}{\partial \zeta_i^2}(\mathcal{K}_i\otimes I_n)G\right). \quad (25)
\end{aligned}
$$

Substituting (25) into (24) gives

$$
\begin{aligned}
\mathscr{L}V_i^*(\zeta_i,t) &= -\zeta_i^T Q_{ii}\zeta_i - \bar{u}_i^2 \bar{R}_{ii}\ln(\underline{\mathbf{1}}_n - \tanh^2(D_i^*)) \\
&\quad + V_{\zeta_i}^{*T}((l_{ii}+b_{ii})\otimes I_n)u_i^*(t_k^i). \quad (26)
\end{aligned}
$$

By (18), one has

$$
\begin{aligned}
&\bar{u}_i^2 \bar{R}_{ii}\ln(\underline{\mathbf{1}} - \tanh^2(D_i^*)) \\
&= M_i(u_i^*) - \bar{u}_i V_{\zeta_i}^{*T}((l_{ii}+b_{ii})\otimes I_n)\tanh(D_i^*). \quad (27)
\end{aligned}
$$

From the definition of $M_i(u_i^*)$, one gets

$$
M_i(u_i^*)=M_i(u_i^*(t_k^i))+2\int_{u_i^*(t_k^i)}^{u_i^*(t)} \bar{u}_i \tanh^{-T}(v_i/\bar{u}_i)R_{ii}dv_i. \quad (28)
$$

Combining (27) and (28) yields

$$
\begin{aligned}
&\bar{u}_i^2 \bar{R}_{ii}\ln(\underline{\mathbf{1}}_n - \tanh^2(D_i^*)) \\
&\quad = M_i(u_i^*(t_k^i)) + 2\int_{u_i^*(t_k^i)}^{u_i^*(t)} \bar{u}_i \tanh^{-T}(v_i/\bar{u}_i)R_{ii}dv_i \\
&\quad\quad - \bar{u}_i V_{\zeta_i}^{*T}((l_{ii}+b_{ii})\otimes I_n)\tanh(D_i^*). \quad (29)
\end{aligned}
$$

Furthermore, owing to $2\bar{u}_i D_i^{*T}R_{ii} = V_{\zeta_i}^{*T}((l_{ii}+b_{ii})\otimes I_n)$, one has

$$
\begin{aligned}
&V_{\zeta_i}^{*T}((l_{ii}+b_{ii})\otimes I_n)u_i^*(t_k^i) \\
&\quad = \int_{u_i^*(t)}^{u_i^*(t_k^i)} V_{\zeta_i}^{*T}((l_{ii}+b_{ii})\otimes I_n)dv_i \\
&\quad\quad + V_{\zeta_i}^{*T}((l_{ii}+b_{ii})\otimes I_n)u_i^*(t) \\
&\quad = \int_{u_i^*(t)}^{u_i^*(t_k^i)} 2\bar{u}_i D_i^{*T}R_{ii}dv_i \\
&\quad\quad - \bar{u}_i V_{\zeta_i}^{*T}((l_{ii}+b_{ii})\otimes I_n)\tanh(D_i^*). \quad (30)
\end{aligned}
$$

Substituting (29) and (30) into (26), one gets

$$
\begin{aligned}
\mathscr{L}V_i^*(\zeta_i,t) &= -\zeta_i^T Q_{ii}\zeta_i - M_i(u_i^*(t_k^i)) \\
&\quad + \int_{u_i^*(t)}^{u_i^*(t_k^i)} 2\bar{u}_i(\tanh^{-1}(v_i/\bar{u}_i) + D_i^*)^T R_{ii}dv_i. \quad (31)
\end{aligned}
$$

Let $v_i=-\bar{u}_i\tanh(\tau_i)$, then $dv_i=-\bar{u}_i\left(\underline{\mathbf{1}}_n - \tanh^2(\tau_i)\right)d\tau_i$. Based on Assumption 3, (31) becomes

$$
\begin{aligned}
\mathscr{L}V_i^*(\zeta_i,t) &= -\zeta_i^T Q_{ii}\zeta_i - M_i(u_i^*(t_k^i)) \\
&\quad - \int_{D_i^*(t)}^{D_i^*(t_k^i)} 2\bar{u}_i^2(-\tau_i + D_i^*)^T(\underline{\mathbf{1}}_n - \tanh^2(\tau_i))R_{ii}d\tau_i \\
&\le -\zeta_i^T Q_{ii}\zeta_i - M_i(u_i^*(t_k^i)) + \int_{D_i^*(t)}^{D_i^*(t_k^i)} 2\bar{u}_i^2(\tau_i - D_i^*)^T R_{ii}d\tau_i \\
&= -\zeta_i^T Q_{ii}\zeta_i - M_i(u_i^*(t_k^i)) \\
&\quad + \bar{u}_i^2\left(D_i^*(\zeta_i(t_k^i)) - D_i^*(\zeta_i)\right)^T R_{ii}(D_i^*(\zeta_i(t_k^i)) - D_i^*(\zeta_i)) \\
&\le -\chi^2\lambda_{\min}(Q_{ii})\|\zeta_i\|^2 - (1-\chi^2)\lambda_{\min}(Q_{ii})\|\zeta_i\|^2 \\
&\quad + \bar{u}_i^2 L_{D_i}^2\|R_{ii}\|\|r_i\|^2. \quad (32)
\end{aligned}
$$

Taking the mathematical expectation on both sides of (32), we can get

$$
\begin{aligned}
\mathbb{E}[\mathscr{L}V_i^*(\zeta_i,t)] &\le -\chi^2\lambda_{\min}(Q_{ii})\mathbb{E}\left[\|\zeta_i\|^2\right] \\
&\quad - \left(1-\chi^2\right)\lambda_{\min}(Q_{ii})\mathbb{E}\left[\|\zeta_i\|^2\right] \\
&\quad + \bar{u}_i^2 L_{D_i}^2\|R_{ii}\|\mathbb{E}\left[\|r_i\|^2\right]. \quad (33)
\end{aligned}
$$

From condition (22) and Assumption 4, one gets

$$
\begin{aligned}
\mathbb{E}[\mathscr{L}V_i^*(\zeta_i,t)] &\le -\chi^2\lambda_{\min}(Q_{ii})\mathbb{E}\left[\|\zeta_i\|^2\right] + e^{-\epsilon t} \\
&\le -c_3\mathbb{E}[V_i^*(\zeta_i,t)] + e^{-\epsilon t} \quad (34)
\end{aligned}
$$

where $c_3=\frac{\chi^2\lambda_{\min}(Q_{ii})}{c_2}>0$. By (23) and (34), we have

$$
\begin{aligned}
&\mathbb{E}[V_i^*(\zeta_i,t)] - \mathbb{E}[V_i^*(z_i,s)] \\
&\quad = \int_s^t \mathbb{E}[\mathscr{L}V_i^*(\zeta_i(\eta),\eta)]d\eta \\
&\quad \le -c_3\int_s^t \mathbb{E}[V_i^*(\zeta_i(\eta),\eta)]d\eta + \int_s^t e^{-\epsilon\eta}d\eta. \quad (35)
\end{aligned}
$$

Thus, we get

$$
\begin{aligned}
\mathbb{E}[V_i^*(\zeta_i,t)] &\le \mathbb{E}[V_i^*(z_i,s)]e^{-c_3(t-s)} + e^{-c_3 t}\int_s^t e^{(c_3-\epsilon)\eta}d\eta \\
&= \mathbb{E}[V_i^*(z_i,s)]e^{-c_3(t-s)} + \frac{1}{c_3-\epsilon}e^{-\epsilon t} \\
&\quad - \frac{1}{c_3-\epsilon}e^{-\epsilon s}e^{-c_3(t-s)} \\
&= \mathbb{E}[V_i^*(z_i,s)]e^{-c_3(t-s)} + e^{-\epsilon s}\varphi_0(t) \quad (36)
\end{aligned}
$$

where $\varphi_0(t) = \frac{1}{c_3-\epsilon}(e^{-\epsilon(t-s)} - e^{-c_3(t-s)}) \ge 0$ and $\lim_{t\to+\infty}\varphi_0(t) = 0$. Therefore, we have $\lim_{t\to+\infty}\mathbb{E}[V_i^*(\zeta_i,t)] = 0$, which means that SMASs (1) achieves the optimal mean-square leader-following consensus eventually under the ETOC strategy.

Step 2. It will prove that the closed loop system does not have Zeno behavior. By (9) and (11), for $t\in[t_k^i, t_{k+1}^i)$, we have

$$
d\zeta_i(t)=(\mathcal{K}_i\otimes I_n)\left(F(e(t),t)dt + u(t)dt + Gdw(t)\right). \quad (37)
$$

Let $\Phi_i(t) = r_i^T(t)r_i(t)$ where $r_i(t) = \zeta_i(t_k^i) - \zeta_i(t)$. Then, from Definition 1, we can obtain that

$$
\mathscr{L}\Phi_i(t) = -2r_i^T(t)[(\mathcal{K}_i\otimes I_n)(F(e(t),t) + u(t))]
$$

$$+\operatorname{Tr}(G^T(\mathcal{K}_i^T\mathcal{K}_i\otimes I_n)G). \quad (38)$$

By Assumption 2, Young's inequality, and $\|G^TG\|\leq Ng_M$, one gets

$$\begin{aligned}&-2r_i^T(t)(\mathcal{K}_i\otimes I_n)F(e(t),t)\\&\leq r_i^T(t)r_i(t)+F^T(e(t),t)(\mathcal{K}_i^T\otimes I_n)(\mathcal{K}_i\otimes I_n)F(e(t),t)\\&\leq r_i^T(t)r_i(t)+L_f^2\zeta_i^T(t)\zeta_i(t)\\&\leq\|r_i(t)\|^2+L_f^2\|\zeta_i(t)\|^2\end{aligned} \quad (39)$$

$$-2r_i^T(t)(\mathcal{K}_i\otimes I_n)u(t)\leq\|r_i(t)\|^2+\|\mathcal{K}_i^T\mathcal{K}_i\|\bar{u}^2 \quad (40)$$

$$\begin{aligned}\operatorname{Tr}(G^T(\mathcal{K}_i^T\mathcal{K}_i\otimes I_n)G)&\leq k\sqrt{k}\|G^T(\mathcal{K}_i^T\mathcal{K}_i\otimes I_n)G\|\\&\leq\frac{k\sqrt{k}}{2}\|\mathcal{K}_i^T\mathcal{K}_i\|^2+\frac{k\sqrt{k}}{2}\|G^TG\|^2\\&\leq\frac{k\sqrt{k}}{2}\|\mathcal{K}_i^T\mathcal{K}_i\|^2+\frac{k\sqrt{k}}{2}N^2g_M^2\end{aligned} \quad (41)$$

where L_f, g_M, and $\bar{u}=\max_{i=1}^N\{\bar{u}_i\}$ are positive constants. Substituting (39), (40) and (41) into (38) yields

$$\begin{aligned}\mathscr{L}\varPhi_i(t)&\leq 2\|r_i(t)\|^2+\|\mathcal{K}_i^T\mathcal{K}_i\|\bar{u}^2\\&\quad+\frac{k\sqrt{k}}{2}\|\mathcal{K}_i^T\mathcal{K}_i\|^2+\frac{k\sqrt{k}}{2}N^2g_M^2+L_f^2\|\zeta_i(t)\|^2\\&=2\|r_i(t)\|^2+\|\mathcal{K}_i^T\mathcal{K}_i\|\bar{u}^2+\frac{k\sqrt{k}}{2}\|\mathcal{K}_i^T\mathcal{K}_i\|^2\\&\quad+\frac{k\sqrt{k}}{2}N^2g_M^2+L_f^2\left(2\|r_i(t)\|^2+2\|\zeta_i(t_k^i)\|^2\right)\\&=a\varPhi_i(t)+b\end{aligned} \quad (42)$$

where $a=2+2L_f^2$ and $b=\|\mathcal{K}_i^T\mathcal{K}_i\|\bar{u}^2+\frac{k\sqrt{k}}{2}\|\mathcal{K}_i^T\mathcal{K}_i\|^2+\frac{k\sqrt{k}}{2}N^2g_M^2+2L_f^2\|\zeta_i(t_k^i)\|^2$ are positive constants. Owing to $r_i\left(t_k^i\right)=0$, during $t\in[t_k^i,t_{k+1}^i)$, we have

$$\mathbb{E}\left[\varPhi_i(t)\right]=\mathbb{E}\left[\|r_i(t)\|^2\right]\leq\frac{\mathbb{E}[b]}{a}\left(e^{a(t-t_k^i)}-1\right). \quad (43)$$

Based on condition (22), when $t=t_{k+1}^i$, we have

$$\begin{aligned}&\mathbb{E}\left[\|r_i(t_{k+1}^i)\|\right]^2\\&=\frac{(1-\chi^2)\lambda_{\min}(Q_{ii})}{\bar{u}_i^2L_{D_i}^2\|R_{ii}\|}\mathbb{E}\left[\|\zeta_i\|^2\right]+\frac{e^{-\epsilon t_{k+1}^i}}{\bar{u}_i^2L_{D_i}^2\|R_{ii}\|}.\end{aligned} \quad (44)$$

Denote $T^i=t_{k+1}^i-t_k^i$. From (43) and (44), when $t=t_{k+1}^i$, we can derive that

$$T^i\geq\frac{1}{a}\ln\left(\frac{a}{\mathbb{E}[b]}\mathbb{E}[\|r_i(t_{k+1}^i)\|^2]+1\right)>0. \quad (45)$$

By (45), we can easily get $T=\inf\{T^i\}>0$ for any i, which means that the closed loop system does not have the Zeno behavior. □

Remark 4 Since event-triggered condition (22) has an exponential term $e^{-\epsilon t}$, it is actually a special dynamic event-triggered strategy. At the same time, the introduction of exponential term $e^{-\epsilon t}$ extends the triggering interval and effectively reduces the number of triggering events.

Proof From (45) and $T^i=t_{k+1}^i-t_k^i>0$, one has

$$\begin{aligned}T^i&\geq\frac{1}{a}\ln\left(\frac{a}{\mathbb{E}[b]}\left(\beta_1\mathbb{E}[\|\zeta_i(t_k^i)\|^2]+\beta_2e^{-\epsilon t_{k+1}^i}\right)+1\right)\\&>\frac{1}{a}\ln\left(\frac{a}{\mathbb{E}[b]}\left(\beta_1\mathbb{E}[\|\zeta_i(t_k^i)\|^2]\right)+1\right)=\bar{T}^i\end{aligned} \quad (46)$$

where $\beta_1=\frac{(1-\chi^2)\lambda_{\min}(Q_{ii})}{\bar{u}_i^2L_{D_i}^2\|R_{ii}\|}$ and $\beta_2=\frac{1}{\bar{u}_i^2L_{D_i}^2\|R_{ii}\|}$. $\bar{T}^i$ is the triggering interval without the exponential term $e^{-\epsilon t}$. Thus, $T=\inf_k\{T^i\}>\inf_k\{\bar{T}^i\}>0$. The proof is completed. □

Remark 5 As we all known that stochastic factors are inevitable in certain dynamical systems, such as the multiple one-link manipulators and the UAVs. Inspired by the optimal control of the constrained-input nonlinear system in Reference [26], this paper considers the ETOC problems for SMASs with input constraints. Moreover, it is also proved that the closed-loop system can exclude the Zeno behavior.

4 Implementation Of approximate optimal controller of SMASs via ET-ACDs

In the process of implementing the approximate ETOC of nonlinear SMASs with unknown internal states and input constraints, it is hard to obtain analytical solutions of stochastic HJB equations via the classical on-policy algorithm. To tackle this issue, the identifier-critic algorithm is utilized where observer-based identifier networks are established to recover the knowledge of unknown system dynamics, and critic networks are established to approximate the value function. The structure of the approximate ETOC algorithm is presented in Fig. 1.

4.1 System identification

According to [41], the neural-network-based observers of SMASs (1) can be designed as follows:

$$\begin{cases}\dot{\hat{x}}_i(t)=A\hat{x}_i(t)+\hat{\omega}_{hi}^T\phi_{hi}(\hat{\mathcal{Z}}_i)+L(y_i-\hat{y}_i)\\\hat{y}_i(t)=C\hat{x}_i(t)\end{cases} \quad (47)$$

where the Hurwitz matrix $A\in\mathbb{R}^{n\times n}$ is chosen to make sure the pair (C,A) observable; $\hat{\mathcal{Z}}_i=Y_{hi}^T[\hat{x}_i^T(t),u_i^T(t)]^T$

with $Y_{hi} \in \mathbb{R}^{(2n\times N_{h,h})}$; $\hat{\omega}_{hi} \in \mathbb{R}^{(N_{h,h}\times n)}$ is an estimation of the ideal weight matrix ω_{hi}; $\phi_{hi} \in \mathbb{R}^{N_{h,h}}$ and $N_{h,h}$ are the activation function and number of the hidden nodes of the identifier network, respectively. Let the state estimation error and identifier weight estimation error be $\tilde{x}_i = x_i - \hat{x}_i$ and $\tilde{\omega}_{hi} = \omega^*_{hi} - \hat{\omega}_{hi}$, respectively. By (1) and (47), the dynamics of state estimation error system become

$$\begin{aligned} d\tilde{x}_i(t) &= A_C\tilde{x}_i(t)dt + (f(x_i(t),t) + u_i(t) - Ax_i(t))dt \\ &\quad - \hat{w}^T_{hi}\phi_{hi}(\hat{\mathcal{Z}}_i)dt + g(x_i(t),t)dw(t) \end{aligned} \tag{48}$$

where $A_C = A - LC$ is a Hurwitz matrix. Based on neural networks, one has

$$f(x_i(t),t)+u_i(t)-Ax_i(t) = \omega^{*T}_{hi}\phi_{hi}(\mathcal{Z}_i)+\varepsilon_{hi}(x_i) \tag{49}$$

where $\mathcal{Z}_i = Y^T_{hi}[x^T_i(t), u^T_i(t)]^T$ and $\varepsilon_{hi} \in \mathbb{R}^n$ is the structural error.

Assumption 5 [35] The ideal identifier weight ω^*_{hi}, activation function ϕ_{hi} and structural error ε_{hi} are bounded by the positive constants ω_{hM}, ϕ_{hM} and ε_{hM} , respectively, i.e., $\|\omega^*_{hi}\|\le\omega_{hM}$, $\|\phi_{hi}\|\le\phi_{hM}$ and $\|\varepsilon_{hi}\|\le\varepsilon_{hM}$.

Theorem 2 *Consider the dynamics of state estimation error system* (48). *Let Assumptions 1-5 valid and the updating rule of the estimated weight matrix* $\hat{\omega}_{hi}$ *is defined as*

$$\dot{\hat{\omega}}_{hi} = -l_{i1}\phi_{hi}\tilde{y}^T_i CA^{-1}_C - l_{i2}\hat{\omega}_{hi} \tag{50}$$

where l_{i1} *and* l_{i2} *are positive weight parameters. Then, the* $\tilde{x}_i$ *and* $\tilde{\omega}_{hi}$ *are SGUUB in mean-square sense.*

Proof Select the Lyapunov function candidate as follows:

$$V_i(\tilde{x}_i(t),t) = \tilde{x}^T_i P\tilde{x}_i + \frac{1}{2}\mathrm{Tr}\left(\tilde{\omega}^T_{hi}l^{-1}_{i1}\tilde{\omega}_{hi}\right) \tag{51}$$

where P is a symmetric positive definite matrix that satisfies $A^T_C P + PA_C = -\beta I_n$ with $\beta > 1$. Let $V_i = V_i(\tilde{x}_i(t),t)$ for simplicity. From Definition 1, we have

$$\begin{aligned} \mathscr{L}V_i &= \tilde{x}^T_i(A^T_C P + PA_C)\tilde{x}_i \\ &\quad + 2\tilde{x}^T_i P\left[(f(x_i(t),t)+u_i(t)-Ax_i(t))-\hat{\omega}^T_{hi}\phi_{hi}(\hat{\mathcal{Z}}_i)\right] \\ &\quad + \mathrm{Tr}\left(g^T Pg\right) + \mathrm{Tr}\left(\tilde{\omega}^T_{hi}l^{-1}_{i1}\dot{\tilde{\omega}}_{hi}\right). \end{aligned} \tag{52}$$

Furthermore, by Assumption 2 and Young's inequality, the following inequalities can be obtained:

$$\begin{aligned} \mathrm{Tr}(g^T Pg) &\le k\sqrt{k}\|g^T Pg\| \le k\sqrt{k}\|P\|\|g^T g\| \\ &\le \frac{k\sqrt{k}}{2}\|P\|^2 + \frac{k\sqrt{k}}{2}g^2_M. \end{aligned} \tag{53}$$

Substituting (49) and (53) into (52) gives

$$\begin{aligned} \mathscr{L}V_i &\le -\beta\|\tilde{x}_i\|^2 \\ &\quad + 2\|\tilde{x}_i\|\|P\|(2\omega_{hM}\phi_{hM} + \varepsilon_{hM} + \|\tilde{\omega}_{hi}\|\phi_{hM}) \\ &\quad + \frac{k\sqrt{k}}{2}\|P\|^2 + \frac{k\sqrt{k}}{2}g^2_M + \mathrm{Tr}\left(\tilde{\omega}^T_{hi}l^{-1}_{i1}\dot{\tilde{\omega}}_{hi}\right). \end{aligned} \tag{54}$$

By (50) and $\mathrm{Tr}(MN) = \mathrm{Tr}(NM) = NM$ for all $M \in \mathbb{R}^{n\times 1}$ and $N \in \mathbb{R}^{1\times n}$, one has

$$\begin{aligned} &\mathrm{Tr}\left(\tilde{\omega}^T_{hi}l^{-1}_{i1}\dot{\tilde{\omega}}_{hi}\right) \\ &\quad = \mathrm{Tr}\left(\tilde{\omega}^T_{hi}l^{-1}_{i1}(l_{i1}\phi_{hi}\tilde{y}^T_i CA^{-1}_C + l_{i2}\hat{\omega}_{hi})\right) \\ &\quad = \tilde{y}^T_i CA^{-1}_C\tilde{\omega}^T_{hi}\phi_{hi} + l_{i2}\mathrm{Tr}\left(\tilde{\omega}^T_{hi}l^{-1}_{i1}\hat{\omega}_{hi}\right). \end{aligned} \tag{55}$$

Based on $\tilde{\omega}_{hi} = \omega^*_{hi} - \hat{\omega}_{hi}$, the following equation can be obtained

$$\begin{aligned} \mathrm{Tr}\left(\tilde{\omega}^T_{hi}l^{-1}_{i1}\hat{\omega}_{hi}\right) &= \frac{1}{2}\mathrm{Tr}\left(\omega^{*T}_{hi}l^{-1}_{i1}\omega^*_{hi}\right) - \frac{1}{2}\mathrm{Tr}\left(\tilde{\omega}^T_{hi}l^{-1}_{i1}\tilde{\omega}_{hi}\right) \\ &\quad - \frac{1}{2}\mathrm{Tr}\left(\hat{\omega}^T_{hi}l^{-1}_{i1}\hat{\omega}_{hi}\right). \end{aligned} \tag{56}$$

Thus, combining (55) and (56), we have

$$\begin{aligned} &\mathrm{Tr}\left(\tilde{\omega}^T_{hi}l^{-1}_{i1}\dot{\tilde{\omega}}_{hi}\right) \\ &\le \tilde{y}^T_i CA^{-1}_C\tilde{\omega}^T_{hi}\phi_{hi} \\ &\quad + l_{i2}\left[\frac{1}{2}\mathrm{Tr}\left(\omega^{*T}_{hi}l^{-1}_{i1}\omega^*_{hi}\right) - \frac{1}{2}\mathrm{Tr}\left(\tilde{\omega}^T_{hi}l^{-1}_{i1}\tilde{\omega}_{hi}\right)\right]. \end{aligned} \tag{57}$$

Substituting (57) into (54) gives

$$\begin{aligned} \mathscr{L}V_i &\le -\beta\|\tilde{x}_i\|^2 + \|\tilde{x}_i\|\psi \\ &\quad + l_{i2}\left[\frac{1}{2}\mathrm{Tr}\left(\omega^{*T}_{hi}l^{-1}_{i1}\omega^*_{hi}\right) - \frac{1}{2}\mathrm{Tr}\left(\tilde{\omega}^T_{hi}l^{-1}_{i1}\tilde{\omega}_{hi}\right)\right] \\ &\quad + \frac{k\sqrt{k}}{2}\|P\|^2 + \frac{k\sqrt{k}}{2}g^2_M \end{aligned} \tag{58}$$

where $\psi = 4\|P\|\omega_{hM}\phi_{hM}+2\|P\|\varepsilon_{hM}+2\|P\|\|\tilde{\omega}_{hi}\|\phi_{hM}+\|C^T CA^{-1}_C\tilde{\omega}^T_{hi}\|\phi_{hM}$. By Young's inequality, one gets

$$\|\tilde{x}_i\|\psi \le \|\tilde{x}_i\|^2 + \frac{1}{4}\psi^2. \tag{59}$$

Substituting (59) into (58) gives

$$\begin{aligned} \mathscr{L}V_i &\le -\frac{(\beta-1)}{\lambda_{\max}(P)}\tilde{x}^T_i P\tilde{x}_i \\ &\quad + l_{i2}\left[\frac{1}{2}\mathrm{Tr}\left(\omega^{*T}_{hi}l^{-1}_{i1}\omega^*_{hi}\right) - \frac{1}{2}\mathrm{Tr}\left(\tilde{\omega}^T_{hi}l^{-1}_{i1}\tilde{\omega}_{hi}\right)\right] \\ &\quad + \frac{1}{4}\psi^2 + \frac{k\sqrt{k}}{2}\|P\|^2 + \frac{k\sqrt{k}}{2}g^2_M \\ &\le -\varpi\left[\tilde{x}^T_i P\tilde{x}_i + \frac{1}{2}\mathrm{Tr}\left(\tilde{\omega}^T_{hi}l^{-1}_{i1}\tilde{\omega}_{hi}\right)\right] + \vartheta \\ &= -\varpi V_i + \vartheta \end{aligned} \tag{60}$$

where $\varpi = \min\left(\frac{(\beta-1)}{\lambda_{\max}(P)}, \frac{l_{i2}}{2}\right)$ and $\vartheta = \frac{l_{i2}}{2}\mathrm{Tr}\left(\omega^{*T}_{hi}l^{-1}_{i1}\omega^*_{hi}\right) + \frac{1}{4}\psi^2 + \frac{k\sqrt{k}}{2}\|P\|^2 + \frac{k\sqrt{k}}{2}g^2_M$ are positive constants. Then, by Lemma 1, we have

$$\mathbb{E}[V_i(\tilde{x}_i(t),t)] \le V_i\left(\tilde{x}_i(0),0\right)e^{-\varpi t} + \frac{\vartheta}{\varpi}, \quad \forall t > 0. \tag{61}$$

which means that $\mathbb{E}[V_i(\tilde{x}_i(t), t)]$ is bounded by ϑ/ϖ eventually. Therefore, we can get that $\tilde{x}_i$ and $\tilde{\omega}_{hi}$ are SGUUB in mean-square sense. □

Remark 6 Reference [21] has studied the optimal control strategy for stochastic nonlinear systems based on the framework of ACDs for the first time, in which event-triggered conditions strictly rely on the known system dynamics, and there is no proof that the Zeno behavior has been excluded. Actually, it is difficult to acquire knowledge of the system dynamics in practical applications completely. To overcome the disadvantage of relying on known system dynamics, the observer-based identifier network is established in this paper, and it is also proved that the Zeno behavior does not exist in the closed loop system via the Itô formula. Additionally, to better meet the needs of actual production, the input-saturation problem is also considered in this paper.

Remark 7 After approximating the system dynamics through the observer-based identifier network, approximate dynamics $\hat{f}$ should be utilized in the subsequent analysis. Actually, it can be seen from (49) and Assumption 5 that the use of approximate dynamics $\hat{f}$ hardly introduces any difficulties. At the same time, to avoid confusion in symbols, the approximate dynamics are still represented by f in the following theoretical analysis.

4.2 Critic network design

By the critic network, the optimal value function can be written as:

$$V_i^*(\zeta_i(t)) = \omega_{ci}^T \phi_{ci}\left(Y_{ci}^T \zeta_i(t)\right) + \varepsilon_{ci} \tag{62}$$

where $Y_{ci} \in \mathbb{R}^{n\times N_{ch}}$ and $\omega_{ci} \in \mathbb{R}^{N_{ch}}$ are the ideal weight matrix of the critic network; $\phi_{ci} \in \mathbb{R}^{N_{ch}}$, ε_{ci}, and N_{ch} are the activation function, structural error, and the number of the hidden nodes of critic network, respectively. The derivative of $V_i^*(\zeta_i(t))$ with respect to $\zeta_i(t)$ is:

$$V_{\zeta_i}^* = \nabla\phi_{ci}^T\left(Y_{ci}^T \zeta_i(t)\right)\omega_{ci} + \nabla\varepsilon_{ci}. \tag{63}$$

Then, the ETOC can be described by

$$u_i^*(t_k^i) = -\bar{u}_i \tanh(D_i^*(t_k^i)), \quad t \in [t_k^i, t_{k+1}^i) \tag{64}$$

where $D_i^*(t_k^i) = \frac{1}{2\bar{u}_i} R_{ii}^{-1}((l_{ii}+b_{ii})\otimes I_n)(\nabla\phi_{ci}^T\omega_{ci} + \nabla\varepsilon_{ci})$.

Assumption 6 [23] In the critic network: (1) ω_{ci}, $\nabla\phi_{ci}$ and $\nabla\varepsilon_{ci}$ are bounded, i.e., $\|\omega_{ci}\| \le \omega_{cM}$, $\|\nabla\phi_{ci}\| \le \nabla\phi_{cM}$, $\|\nabla\varepsilon_{ci}\| \le \nabla\varepsilon_{cM}$ where ω_{cM}, $\nabla\phi_{cM}$, and $\nabla\varepsilon_{cM}$ are positive constants.

Substituting (63) and (64) into (21), the event-triggered stochastic HJB equation becomes

$$\begin{aligned} &V_{i,t}^* + H_i(\zeta_i, \omega_{ci}, u_i^*(t_k^i), u_{(i)}^*, t) \\ &\quad = V_{i,t}^* + \zeta_i^T Q_{ii}\zeta_i + M_i(u_i^*(t_k^i)) \\ &\qquad + \omega_{ci}^T \nabla\phi_{ci}((l_{ii}+b_{ii})\otimes I_n)(\tilde{f}_i - \bar{u}_i\tanh(D_i^*(t_k^i))) \\ &\qquad + \omega_{ci}^T\nabla\phi_{ci}\sum_{j\in\mathcal{N}_i}((l_{ij}+b_{ij})\otimes I_n)(\tilde{f}_j - \bar{u}_j\tanh(D_j^*)) \\ &\qquad + \frac{1}{2}\mathrm{Tr}\left(G^T(\mathcal{K}_i^T\otimes I_n)\omega_{ci}^T\frac{\partial^2\phi_{ci}}{\partial\zeta_i^2}(\mathcal{K}_i\otimes I_n)G\right) - \varepsilon_{Hi} \\ &\quad = \varepsilon_{ui}, \qquad\qquad t\in[t_i^k, t_{i+1}^k) \end{aligned} \tag{65}$$

where $\varepsilon_{Hi} = -\nabla\varepsilon_{ci}^T(\mathcal{K}_i\otimes I_n)(F(e(t),t)+u(t))$ is a bounded residual error, i.e., $|\varepsilon_{Hi}| \le \varepsilon_{HM}$ with $\varepsilon_{HM} > 0$. Let the weight estimation error be $\tilde{\omega}_{ci} = \omega_{ci} - \hat{\omega}_{ci}$ with $\hat{\omega}_{ci}$ being an estimation of ω_{ci}. Then, the actual value function can be approximated by

$$\hat{V}_i(\zeta_i) = \hat{\omega}_{ci}^T\phi_{ci}\left(Y_{ci}^T\zeta_i(t)\right) \tag{66}$$

where $\hat{V}_i(\zeta_i)$ is an estimation of $V_i^*(\zeta_i(t))$. Thus, the approximate ETOC is described by

$$\hat{u}_i^*(t_k^i) = -\bar{u}_i\tanh(\hat{D}_i^*(t_k^i)), \quad t\in[t_k^i, t_{k+1}^i) \tag{67}$$

where $\hat{D}_i^*(t_k^i) = \frac{1}{2\bar{u}_i}R_{ii}^{-1}((l_{ii}+b_{ii})\otimes I_n)\nabla\phi_{ci}^T\hat{\omega}_{ci}$. Substituting (64) and (66) into (21), the estimated event-triggered stochastic HJB equation becomes

$$\begin{aligned} &V_{i,t}^* + H_i(\zeta_i, \hat{\omega}_{ci}, u_i^*(t_k^i), u_{(i)}^*, t) \\ &\quad = V_{i,t}^* + \zeta_i^T Q_{ii}\zeta_i + M_i(u_i^*(t_k^i)) \\ &\qquad + \hat{\omega}_{ci}^T \nabla\phi_{ci}((l_{ii}+b_{ii})\otimes I_n)(\tilde{f}_i - \bar{u}_i\tanh(D_i^*(t_k^i))) \\ &\qquad + \hat{\omega}_{ci}^T\nabla\phi_{ci}\sum_{j\in\mathcal{N}_i}((l_{ij}+b_{ij})\otimes I_n)(\tilde{f}_j - \bar{u}_j\tanh(D_j^*)) \\ &\qquad + \frac{1}{2}\mathrm{Tr}\left(G^T(\mathcal{K}_i^T\otimes I_n)\hat{\omega}_{ci}^T\frac{\partial^2\phi_{ci}}{\partial\zeta_i^2}(\mathcal{K}_i\otimes I_n)G\right) \\ &\quad = \varepsilon_i, \qquad\qquad t\in[t_i^k, t_{i+1}^k) \end{aligned} \tag{68}$$

where ε_i is the estimated residual error. By (65) and (68), one gets

$$\begin{aligned} \varepsilon_i = &-\tilde{\omega}_{ci}^T\nabla\phi_{ci}(\mathcal{K}_i\otimes I_n)(F(e(t),t)+u(t)) \\ &+ \varepsilon_{Hi} + \varepsilon_{ui}. \end{aligned} \tag{69}$$

Then, the gradient descent method is used to minimize the following error function E_i

$$E_i = \frac{1}{2}\varepsilon_i^2. \tag{70}$$

Thus, the updated rule of estimated weight $\hat{\omega}_{ci}$ is as follows:

$$\begin{cases} \dot{\hat{\omega}}_{ci} = 0, \ t\in(t_k^i, t_{k+1}^i) \\ \hat{\omega}_{ci}^+ = \hat{\omega}_{ci} - l_{ci}\left(\frac{k_{i1}k_{i1}^T}{(k_{i1}^Tk_{i1}+1)^2}\hat{\omega}_{ci} + \frac{k_{i1}}{(k_{i1}^Tk_{i1}+1)^2}\Psi_i\right), t = t_k^i \end{cases} \tag{71}$$

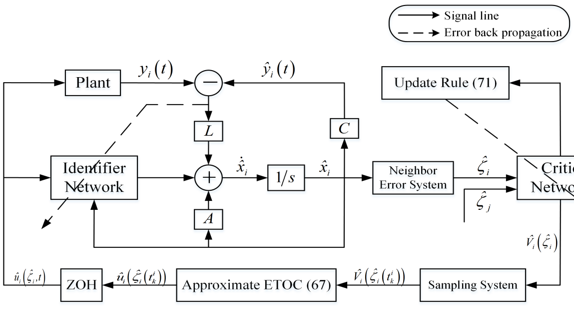

Fig. 1. Diagram of approximate ETOC mechanism

where $k_{i1}=\nabla\phi_{ci}((l_{ii}+b_{ii})\otimes I_n)(\tilde{f}_i-\bar{u}_i\tanh(D_i^*(t_k^i)))+\nabla\phi_{ci}\sum_{j\in\mathcal{N}_i}((l_{ij}+b_{ij})\otimes I_n)(\tilde{f}_j-\bar{u}_j\tanh(D_j^*))\in\mathbb{R}^{N_{ch}}$, $\Psi_i=V_{i,t}^*+\zeta_i^TQ_{ii}\zeta_i+M_i(u_i^*(t_k^i))+\frac{1}{2}\mathrm{Tr}(G^T(\mathcal{K}_i^T\otimes I_n)\times\omega_{ci}^T\frac{\partial^2\phi_{ci}}{\partial\zeta_i^2}(\mathcal{K}_i\otimes I_n)G)$ and $l_{ci}>0$ is the learning rate.

Remark 8 From (21), one has $\varepsilon_{ui}=H_i(\zeta_i,V_{\zeta_i}^*,u_i^*(t_k^i),u_{(i)}^*,t)-H_i(\zeta_i,V_{\zeta_i}^*,u_i^*,u_{(i)}^*,t)=M_i(u_i^*(t_k^i))-M_i(u_i^*)+2\bar{u}_iD_i^{*T}\times R_{ii}(u_i^*(t_k^i)-u_i^*)$. Firstly, we know from (18) and Remark 2 that $M_i(u_i^*(t_k^i))$ and $M_i(u_i^*)$ are bounded. Then, by $\|\omega_{ci}\|\le\omega_{cM}$, $\|\nabla\phi_{ci}\|\le\nabla\phi_{cM}$, and $\|\nabla\varepsilon_{ci}\|\le\nabla\varepsilon_{cM}$ in Assumption 6, one gets $2D_i^{*T}\bar{u}_iR_{ii}(u_i^*(t_k^i)-u_i^*)$ is also bounded. Thus, the transformation error ε_{ui} is bounded, i.e., $|\varepsilon_{ui}|\le\varepsilon_{uM}$ with $\varepsilon_{uM}>0$.

Remark 9 ε_{Hi} is a residual error caused by structural error ε_{ci} in the critic network. Theoretically, the ideal weight vector ω_{ci} is updated along the gradient direction where the residual error ε_{Hi} tends to zero. However, since ω_{ci} is unknown, the estimated weight $\hat{\omega}_{ci}$ is used to approximate the actual output $\hat{V}_i$. In this process, due to the weight estimation error $\tilde{\omega}_{ci}$, the estimated residual error ε_i is generated. What's more, $\varepsilon_i=-\tilde{\omega}_{ci}^T\nabla\phi_{ci}(\mathcal{K}_i\otimes I_n)(F(e(t),t)+u(t))+\varepsilon_{Hi}+\varepsilon_{ui}$. Thus, the estimated weight $\hat{\omega}_{ci}$ is updated along the gradient direction where the estimated residual error ε_i tends to zero. When $\varepsilon_{ci}\to 0$, the approximate optimal value function approaches to the optimal value function eventually, i.e.,$\hat{V}_i(\zeta_i)\to V_i^*(\zeta_i)$. By formula (67), we can get that $\hat{u}_i^*(t_k^i)\to u_i^*(t_k^i)$.

Theorem 3 *Consider the nonlinear constrained-input SMASs* (1) *with event-triggered strategy* (67). *The weights of the critic network are updated as* (71). *Let Assumptions 1-6 valid and the event-triggered condition satisfies*

$$\mathbb{E}\left[\|r_i(t)\|^2\right]\le\frac{(1-\chi^2)\lambda_{\min}(Q_{ii})}{a_1}\mathbb{E}\left[\|\zeta_i\|^2\right]+\frac{e^{-\epsilon t}}{a_1}\quad(72)$$

where $a_1=2\bar{u}_i^2L_{D_i}^2\|R_{ii}\|$ and $0<\chi<1$ are positive constants. Then, $\zeta_i(t)$ and $\tilde{\omega}_{ci}$ are SGUUB in mean-square sense and the SMASs (1) *will achieve the optimal mean-square consensus eventually. Moreover, it is also proved that the closed-loop system can exclude the Zeno behavior.*

Proof Step 1. The following two cases are considered.

Case 1: Events are not triggered ($t_k^i<t<t_{k+1}^i$). Construct a candidate Lyapunov function as follows:

$$V_i(t)=V_{i1}(t)+V_{i2}(t)+V_{i3}(t)\quad(73)$$

where $V_{i1}(t)=V_i^*(\zeta_i(t),t)$, $V_{i2}(t)=V_i^*(\zeta_i(t_k^i),t_k^i)$, $V_{i3}(t)=\frac{1}{l_{ci}}\mathrm{Tr}(\tilde{\omega}_{ci}^T\tilde{\omega}_{ci})$. During $t\in(t_k^i,t_{k+1}^i)$, $\mathscr{L}V_{i2}(t)=\mathscr{L}V_{i3}(t)=0$. Based on Theorem 1, we have

$$\begin{aligned}\mathscr{L}V_{i1}^*(t)&=\mathscr{L}V_i^*(\zeta_i,t)\\&\le-\zeta_i^TQ_{ii}\zeta_i-M_i(\hat{u}_i^*(t_k^i))+\int_{D_i^*(t)}^{\hat{D}_i^*(t_k^i)}2\bar{u}_i^2(\tau_i-D_i^*)^TR_{ii}d\tau_i\\&=-\zeta_i^TQ_{ii}\zeta_i+\bar{u}_i^2(\hat{D}_i^*(t_k^i)-D_i^*(t))^TR_{ii}(\hat{D}_i^*(t_k^i)-D_i^*(t))\\&\le-\chi^2\lambda_{\min}(Q_{ii})\|\zeta_i\|^2-(1-\chi^2)\lambda_{\min}(Q_{ii})\|\zeta_i\|^2\\&\quad+\bar{u}_i^2\|R_{ii}\|\left\|\hat{D}_i^*(t_k^i)-D_i^*(t)\right\|^2.\end{aligned}\quad(74)$$

According to Assumption 3 and 6, one gets

$$\begin{aligned}&\|\hat{D}_i^*(t_k^i)-D_i^*(t)\|\\&\le\|\hat{D}_i^*(t_k^i)-D_i^*(t_k^i)\|+\|D_i^*(t_k^i)-D_i^*(t)\|\\&\le L_R\|\nabla\phi_{ci}^T\hat{\omega}_{ci}-(\nabla\phi_{ci}^T\omega_{ci}+\nabla\varepsilon_{ci})\|+L_{D_i}\|r_i\|\\&\le L_R(\nabla\phi_{cM}\|\tilde{\omega}_{ci}\|+\nabla\varepsilon_{cM})+L_{D_i}\|r_i\|.\end{aligned}\quad(75)$$

Substituting (75) into (74) and taking the expectation yields

$$\begin{aligned}\mathbb{E}[\mathscr{L}V_i^*(\zeta_i,t)]\le&-\chi^2\lambda_{\min}(Q_{ii})\mathbb{E}\left[\|\zeta_i\|^2\right]\\&-(1-\chi^2)\lambda_{\min}(Q_{ii})\mathbb{E}\left[\|\zeta_i\|^2\right]\\&+a_1\mathbb{E}\left[\|r_i(t)\|^2\right]+a_2\end{aligned}\quad(76)$$

where $a_1=2\bar{u}_i^2L_{D_i}^2\|R_{ii}\|$, $a_2=2\bar{u}_i^2L_R^2\|R_{ii}\|(\nabla\phi_{cM}\|\tilde{\omega}_{ci}\|+\nabla\varepsilon_{cM})^2$, and $L_R=\|\frac{1}{2\bar{u}_i}R_{ii}^{-1}((l_{ii}+b_{ii})\otimes I_n)\|$. According to event-triggered condition (72), we have

$$\begin{aligned}\mathbb{E}[\mathscr{L}V_i^*(\zeta_i,t)]&\le-\chi^2\lambda_{\min}(Q_{ii})\mathbb{E}\left[\|\zeta_i\|^2\right]+a_2+e^{-\epsilon t}\\&\le-a_0\mathbb{E}[V_i^*(\zeta_i,t)]+a_2+e^{-\epsilon t}\end{aligned}\quad(77)$$

where $a_0=\frac{\chi^2\lambda_{\min}(Q_{ii})}{c_2}>0$. Similar to (34)-(36), one has

$$\begin{aligned}\mathbb{E}[V_i^*(\zeta_i,t)]\le&\mathbb{E}[V_i^*(z_i,s)]e^{-a_0(t-s)}+\frac{1}{a_0-\epsilon}e^{-\epsilon t}\\&-\frac{1}{a_0-\epsilon}e^{-\epsilon s}e^{-a_0(t-s)}+\frac{a_2}{a_0}\\=&\mathbb{E}[V_i^*(z_i,s)]e^{-a_0(t-s)}+e^{-\epsilon s}\varphi_1(t)+\frac{a_2}{a_0}\end{aligned}\quad(78)$$

where $\varphi_1(t) = \frac{1}{a_0-\epsilon}(e^{-\epsilon(t-s)} - e^{-a_0(t-s)}) \geq 0$ and $\lim_{t\to+\infty} \varphi_1(t) = 0$. Thus, $\zeta_i(t)$ and $\tilde{w}_{ci}$ are SGUUB in mean-square sense. Furthermore, we can conclude that $\lim_{t\to+\infty} \mathbb{E}[V_i^*(\zeta_i, t)] = 0$ by choosing a proper parameter a_2 small enough.

Case 2: Events are triggered ($t = t_k^i$). For the same Lyapunov function:

$$V_i(t) = V_{i1}(t) + V_{i2}(t) + V_{i3}(t). \tag{79}$$

Since $\hat{u}_i^*(t_k^i)$ is an admissible control on $\mathcal{U}_i[s, +\infty)$, one gets $\mathbb{E}[V_i^*(\zeta_i^+(t), t)]$, $\mathbb{E}[V_i^*(\zeta_i(t), t)]$, $\mathbb{E}[V_i^*(\zeta_i^+(t_k^i), t_k^i)]$, and $\mathbb{E}[V_i^*(\zeta(t_k^i), t_k^i)]$ are finite, i.e., $\Phi_{Li} \leq \mathbb{E}[V_i^*] \leq \Phi_{Ui}$ where Φ_{Li} and Φ_{Ui} are positive constants. Thus,

$$\mathbb{E}[\Delta V_{i1}(t)] = \mathbb{E}[V_i^*(\zeta_i^+(t), t)] - \mathbb{E}[V_i^*(\zeta_i(t), t)] \leq 2\Phi_{Ui} \tag{80}$$

$$\mathbb{E}[\Delta V_{i2}(t)] = \mathbb{E}[V_i^*(\zeta_i^+(t_k^i), t_k^i)] - \mathbb{E}[V_i^*(\zeta_i(t_k^i), t_k^i)] \leq 2\Phi_{Ui}. \tag{81}$$

The first-order difference of the third term is: $\Delta V_{i3}(t) = 1/l_{ci} \cdot (\mathrm{Tr}(\tilde{\omega}_{ci}^{T+}\tilde{\omega}_{ci}^{+}) - \mathrm{Tr}(\tilde{\omega}_{ci}^{T}\tilde{\omega}_{ci}))$. Based on $\hat{\omega}_{ci}^{+} = \hat{\omega}_{ci} - l_{ci}(\frac{k_{i1}k_{i1}^T}{(k_{i1}^T k_{i1}+1)^2}\hat{\omega}_{ci} + \frac{k_{i1}}{(k_{i1}^T k_{i1}+1)^2}\Psi_i)$ and (65), we have

$$\begin{aligned} \tilde{\omega}_{ci}^{+} &= \tilde{\omega}_{ci} - l_{ci}(\Xi_1\tilde{\omega}_{ci} - \Xi_2(k_{i1}^T\omega_{ci} + \Psi_i)) \\ &= \tilde{\omega}_{ci} - l_{ci}(\Xi_1\tilde{\omega}_{ci} - \Xi_2\varepsilon_{Ti}) \end{aligned} \tag{82}$$

where $\Xi_1 = \frac{k_{i1}k_{i1}^T}{(k_{i1}^T k_{i1}+1)^2} \in \mathbb{R}^{N_{ch}\times N_{ch}}$ and $\Xi_2 = \frac{k_{i1}}{(k_{i1}^T k_{i1}+1)^2} \in \mathbb{R}^{N_{ch}}$ are utilized for normalization, $\varepsilon_{Ti} = \varepsilon_{Hi} + \varepsilon_{ui}$ with $|\varepsilon_{Ti}| \leq \varepsilon_{HM} + \varepsilon_{uM} = \varepsilon_{TM}$. Then,

$$\begin{aligned} &\mathrm{Tr}(\tilde{\omega}_{ci}^{T+}\tilde{\omega}_{ci}^{+}) - \mathrm{Tr}(\tilde{\omega}_{ci}^{T}\tilde{\omega}_{ci}) \\ &= \mathrm{Tr}\left((\tilde{\omega}_{ci} - l_{ci}(\Xi_1\tilde{\omega}_{ci} - \Xi_2\varepsilon_{Ti}))^T(\tilde{\omega}_{ci} - l_{ci}(\Xi_1\tilde{\omega}_{ci} - \Xi_2\varepsilon_{Ti}))\right) \\ &\quad - \mathrm{Tr}(\tilde{\omega}_{ci}^{T}\tilde{\omega}_{ci}) \\ &= -2l_{ci}\mathrm{Tr}\left(\tilde{\omega}_{ci}^{T}(\Xi_1\tilde{\omega}_{ci} - \Xi_2\varepsilon_{Ti})\right) \\ &\quad + l_{ci}^2\mathrm{Tr}\left((\Xi_1\tilde{\omega}_{ci} - \Xi_2\varepsilon_{Ti})^T(\Xi_1\tilde{\omega}_{ci} - \Xi_2\varepsilon_{Ti})\right). \end{aligned} \tag{83}$$

By the definition of Ξ_1 and Ξ_2, we can derive that $\|\Xi_1\| \leq \frac{1}{4}$ and $\|\Xi_2\| \leq \frac{3\sqrt{3}}{16}$. Thus,

$$\begin{aligned} &\mathrm{Tr}(\tilde{\omega}_{ci}^{T+}\tilde{\omega}_{ci}^{+}) - \mathrm{Tr}(\tilde{\omega}_{ci}^{T}\tilde{\omega}_{ci}) \\ &\quad = -2l_{ci}\tilde{\omega}_{ci}^{T}\Xi_1\tilde{\omega}_{ci} + 2l_{ci}\tilde{\omega}_{ci}^{T}\Xi_2\varepsilon_{Ti} \\ &\qquad + l_{ci}^2\mathrm{Tr}\left((\Xi_1\tilde{\omega}_{ci} - \Xi_2\varepsilon_{Ti})^T(\Xi_1\tilde{\omega}_{ci} - \Xi_2\varepsilon_{Ti})\right) \\ &\quad \leq -2l_{ci}\lambda_{\min}(\Xi_1)\|\tilde{\omega}_{ci}\|^2 + \frac{3\sqrt{3}}{8}l_{ci}\|\tilde{\omega}_{ci}\|\varepsilon_{TM} \\ &\qquad + l_{ci}^2\|\Xi_1\tilde{\omega}_{ci} - \Xi_2\varepsilon_{Ti}\|^2. \end{aligned} \tag{84}$$

According Young's Inequality, we can obtain

$$l_{ci}\|\tilde{\omega}_{ci}\|\varepsilon_{TM} \leq \frac{l_{ci}}{2}\left(\|\tilde{\omega}_{ci}\|^2 + \varepsilon_{TM}^2\right) \tag{85}$$

$$\begin{aligned} l_{ci}^2\|\Xi_1\tilde{\omega}_{ci} - \Xi_2\varepsilon_{Ti}\|^2 &\leq l_{ci}^2(2\Xi_1^2\|\tilde{\omega}_{ci}\|^2 + 2\Xi_2^2\varepsilon_{TM}^2) \\ &\leq l_{ci}^2(\frac{1}{8}\|\tilde{\omega}_{ci}\|^2 + \frac{27}{128}\varepsilon_{TM}^2). \end{aligned} \tag{86}$$

Substituting (85) and (86) into (84), one gets

$$\begin{aligned} &\mathrm{Tr}(\tilde{\omega}_{ci}^{T+}\tilde{\omega}_{ci}^{+}) - \mathrm{Tr}(\tilde{\omega}_{ci}^{T}\tilde{\omega}_{ci}) \\ &\quad \leq -2l_{ci}\lambda_{\min}(\Xi_1)\|\tilde{\omega}_{ci}\|^2 + \frac{3\sqrt{3}}{16}l_{ci}(\|\tilde{\omega}_{ci}\|^2 + \varepsilon_{TM}^2) \\ &\qquad + l_{ci}^2(\frac{1}{8}\|\tilde{\omega}_{ci}\|^2 + \frac{27}{128}\varepsilon_{TM}^2) \\ &\quad \leq (-2l_{ci}\lambda_{\min}(\Xi_1) + \frac{3\sqrt{3}}{16}l_{ci} + \frac{l_{ci}^2}{8})\|\tilde{\omega}_{ci}\|^2 \\ &\qquad + (\frac{3\sqrt{3}}{16}l_{ci} + \frac{27}{128}l_{ci}^2)\varepsilon_{TM}^2. \end{aligned} \tag{87}$$

Thus, the first-order difference of $V_{i1}(t)$ is:

$$\begin{aligned} \Delta V_{i3}(t) &= (\mathrm{Tr}(\tilde{\omega}_{ci}^{T+}\tilde{\omega}_{ci}^{+}) - \mathrm{Tr}(\tilde{\omega}_{ci}^{T}\tilde{\omega}_{ci}))/l_{ci} \\ &\leq -a_3\|\tilde{\omega}_{ci}\|^2 + a_4 \end{aligned} \tag{88}$$

where $a_3 = 2\lambda_{\min}(\Xi_1) - \frac{3\sqrt{3}}{16} - \frac{l_{ci}}{8}$ and $a_4 = (\frac{3\sqrt{3}}{16} + \frac{27}{128}l_{ci})\varepsilon_{TM}^2$. Thus, based on (80), (81) and (88), one has $\mathbb{E}[\Delta V_i(t)] \leq 0$ when $\|\tilde{\omega}_{ci}\| > ((a_4 + 4\Phi_{Ui})/a_3)^{1/2}$. That is, $\mathbb{E}[V_i(\zeta_i(t_k^i))] \leq \mathbb{E}[V_i(\zeta_i(s))]$, i.e., all error signals are SGUUB at the triggering instants.

From Cases 1 and 2, we can conclude that $\zeta_i(t)$ and $\tilde{\omega}_{ci}$ are SGUUB in mean-square sense and the SMASs (1) will achieve the optimal mean-square leader-following consensus under ETC mechanism eventually.

Step 2. Let $\Phi_i(t) = r_i^T(t)r_i(t)$ where $r_i(t) = \zeta_i(t_k^i) - \zeta_i(t)$. Based on (67), identifier-critic networks, and Assumptions 1-6, it's easy to derive that $\mathbb{E}[\mathscr{L}\Phi_i(t)] \leq \bar{a}\mathbb{E}[\Phi_i(t)] + \mathbb{E}[\bar{b}]$ during $t \in [t_k^i, t_{k+1}^i)$ where $\bar{a}$ and $\mathbb{E}[\bar{b}]$ are positive constants. Then, the rest of the proofs for the exclusion of Zeno behavior are similar to Step 2 of Theorem 1, which are omitted here. Thus, the closed-loop system under event-triggered condition (72) can exclude the Zeno behavior. □

Remark 10 Literatures [35, 42] adopted the actor-critic algorithm to design the approximate optimal control where the boundedness of the weights of actor network and critic network both need to be proved. To simplify the architecture of the algorithm, Zhang et al. [11] proposed a critic-only network to solve the optimal consensus problem of MASs by using the relationship between optimal control and value function induced by the HJB equation, which eliminates the approximation error of the actor network and improves the efficiency of weight updating. Thus, by using the critic-only network, the ETOC strategies for SMASs with input constraints are designed in this paper.

5 Simulation

This section is devoted to verifying the validity of the theoretical analysis using two examples.

Example 1 Consider SMASs (1) with the directed communication topology shown in Fig. 2. The edge weights

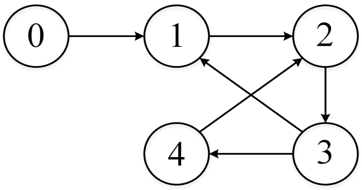

Fig. 2. Directed communication topology

of all followers are defined as $a_{13} = a_{21} = a_{24} = a_{32} = a_{43} = 1$ and the pinning matrix is $\mathcal{B} = \mathrm{diag}\{1, 0, 0, 0\}$. Let initial states of agent i and the leader be $x_i(0) = [i, 3i]^T, (i = 1, 2, 3, 4)$ and $x_0(0) = [2, 1]^T$, respectively. Similar to [19], the dynamics of agents are selected as: $f(x_i) = [-x_{i1} + x_{i2}; -0.5\,(x_{i1} + x_{i2}) - 0.5x_{i2}\sin^2(x_{i1})]$ and $g(x_i) = [2 + \sin^2(x_{i1}); 1 + \cos^2(x_{i2})]$ with $x_i = [x_{i1}; x_{i2}]$. Let the step size of each update be 0.01s.

Firstly, the observer-based identifier network is utilized to recover the knowledge of unknown system dynamics instead of utilizing the system dynamics directly. The structure of the identifier network is 4-6-2 (four input neurons, six hidden neurons and two output neurons) and the matrices in (47) are selected as $A = [-1, 3; -3.5, -5]$ and $L = [1; -0.1]$. Secondly, the structure of the critic network is 2-3-1 where $\phi_{ci}(\zeta_i) = [\tanh(\zeta_{i1}^2); \tanh(\zeta_{i1}\zeta_{i2}); \tanh(\zeta_{i2}^2)]$ with $\zeta_i = [\zeta_{i1}; \zeta_{i2}]$ for $i = 1, 2 \cdots, 4$. The event-triggered parameters in condition (72) are selected as: $\bar{u}_i = 1$, $L_{D_i} = 1.83$, $Q_{ii} = 0.01I_2$, $R_{ii} = I_2$, $\chi = 0.58$ and $\epsilon = 1$. In addition, the parameters in the identifier-critic algorithm are depicted in Table 1:

Table 1: Parameters of the identifier-critic algorithm

Parameter	Meaning	Value
l_{ci}	Learning rate of CN	1×10^{-3}
l_{i1}, l_{i2}	Learning rate of IN	1×10^{-2}
N_c, N_i	Internal cycles of CN and IN	200
T_{Ec}, T_{Ei}	Thresholds of CN and IN	e^{-10}

Fig. 3 (a)-(b) demonstrate the real and identified state trajectories $x_i(t)$ and $\hat{x}_i(t)$. Fig. 3 (c)-(d) demonstrate two components of the identifier error $\tilde{x}_i(t)$ of all followers. Fig. 4 displays weights of the identifier-critic network. The triggering thresholds and measurement errors of all followers under ET-ACDs are depicted in Fig. 5. Fig. 6 shows the approximate ETOC of all followers under ET-ACDs and CT-ACDs, respectively. Fig. 7 demonstrates the cumulative costs of all followers under ET-ACDs, CT-ACDs, ET, and CT strategies where the traditional control gain is $K = [0.6, 0.5]$. It can be seen that ACDs can effectively save cumulative costs. Fig. 8 shows the event-triggered instants under ET-ACDs.

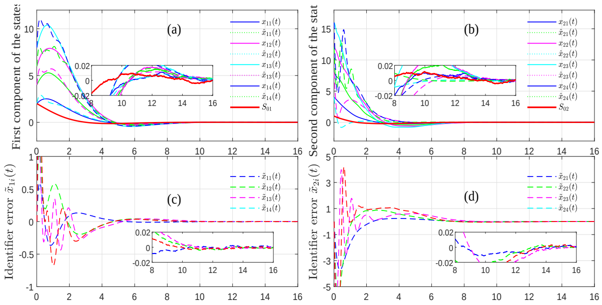

Fig. 3. State components and identifier errors under ET-ACDs

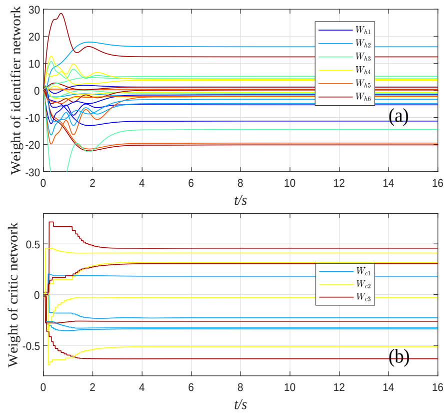

Fig. 4. Network weights. (a) Identifier network (b) Critic network

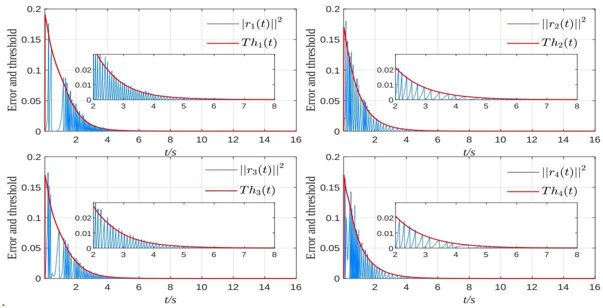

Fig. 5. Triggering thresholds and measurement errors under ET-ACDs

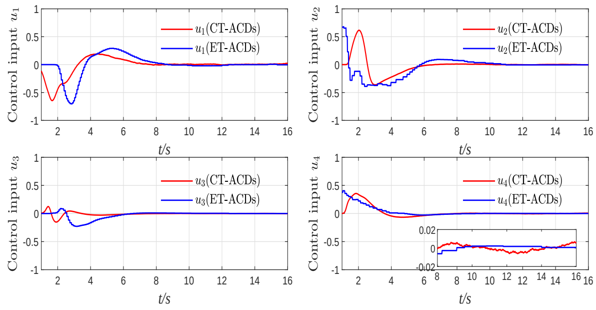

Fig. 6. Approximate ETOC under ET-ACDs and CT-ACDs.

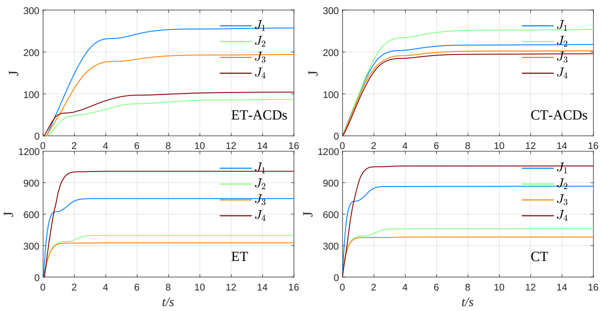

Fig. 7. Cumulative costs under ET-ACDs, CT-ACDs, ET and CT strategies

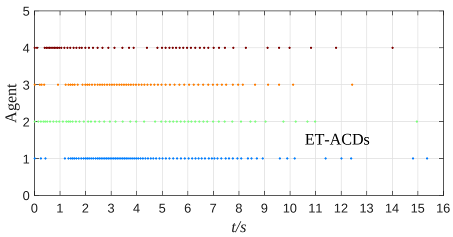

Fig. 8. The event-triggered instants under ET-ACDs

Example 2 During the operation of one-link manipulator, the friction and impact force between mechanical parts will cause random vibration of parts and shells. Therefore, consider four one-link manipulators in random vibration environments with the communication topology shown in Fig. 9.

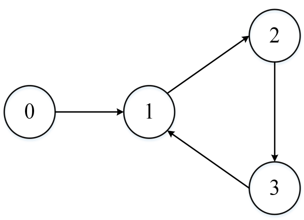

Fig. 9. Directed communication topology.

The dynamics of each manipulator is described by [43]:

$$
\begin{aligned}
dq_{i1} &= q_{i2}dt \\
dq_{i2} &= J_i^{-1}[u_i - B_r^i q_{i2} - M_i g l_i \sin(q_{i1})]dt + g_i \Sigma dw
\end{aligned} \tag{89}
$$

where $q_i = [q_{i1}; q_{i2}]$ and u_i are the state and control of manipulator i. The physical meaning of all the parameters J_i, B_r^i, M_i, g, and l_i can be found in [43]. w represents the 2-dimensional normal Brownian motion to describe the random vibration environment. The values of these parameters are selected as $J_i = 2$, $B_r^i = 0.04$, $M_i = \frac{1}{3}$, $g = 9.8m/s^2$, $l_i = \frac{3}{10}$, $g_i = [0; \sin(q_{i1})]$, $\Sigma = \mathrm{diag}\{0.001, 0.001\}$ respectively. Then, the system (89) can be rewritten as

$$
\begin{aligned}
dq_{i1} &= q_{i2}dt \\
dq_{i2} &= [-0.49\sin(q_{i1}) - 0.02q_{i2} + 0.5u_i]dt + g_i \Sigma dw.
\end{aligned} \tag{90}
$$

The dynamics of the leader is

$$
\begin{aligned}
dq_{01} &= q_{02}dt \\
dq_{02} &= [-0.49\sin(q_{01}) - 0.02q_{02}]dt + g_0 \Sigma dw.
\end{aligned} \tag{91}
$$

To verify the validity of the proposed algorithm, the ET-ACDs and traditional ET strategies are applied to the multiple one-link manipulators, respectively, with the control gain of ET strategy being $K = [2, 3]$. For ET-ACDs, the structure of the identifier network and critic network are 4-6-2 and 2-3-1, respectively, and the activation function is the same as that in Example 1. The control time $T_{total} = 60s$ and the step size of each update is 0.1s.

Fig. 10 (a)-(b) demonstrate the states of all agents with/without ET-ACDs strategies, respectively. Fig. 10 (c)-(d) demonstrate two components of the identifier error. The triggering thresholds and measurement errors of all followers under ET-ACDs is depicted in Fig. 11. Fig. 12 shows the approximate ETOC of all followers under ET-ACDs and ET strategies, respectively. Fig. 13 displays the weights of the identifier-critic network. Fig. 14 presents the event-triggered instants under ET-ACDs with the triggering numbers of all followers being 198, 219, 194, respectively. Fig. 15 demonstrates the cumulative costs of all followers under ET-ACDs and ET strategies, respectively. It can be seen that the framework of ACDs can effectively save the cumulative costs.

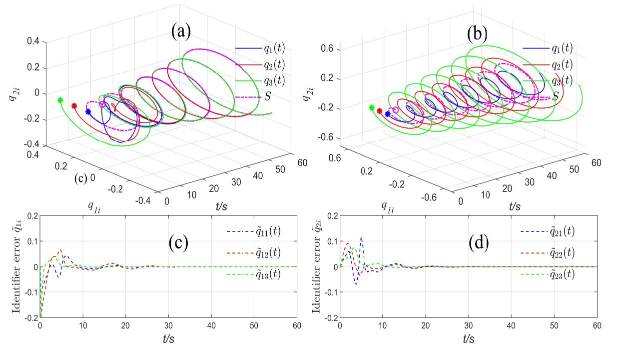

Fig. 10. (a) States under ET-ACDs. (b) States without ET-ACDs. (c) Identifier error $\tilde{q}_{1i}$. (d) Identifier error $\tilde{q}_{2i}$

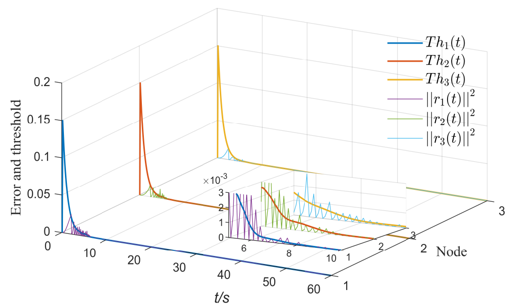

Fig. 11. Triggering thresholds and measurement errors under ET-ACDs

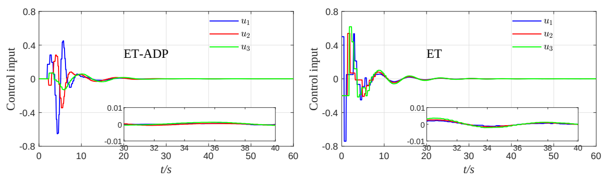

Fig. 12. Control strategies under ET-ACDs and ET strategies

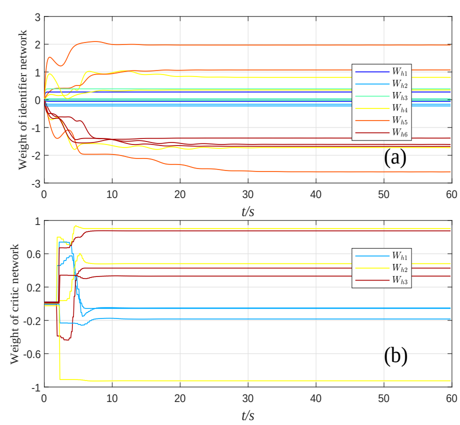

Fig. 13. Network weights. (a) Identifier network (b) Critic network

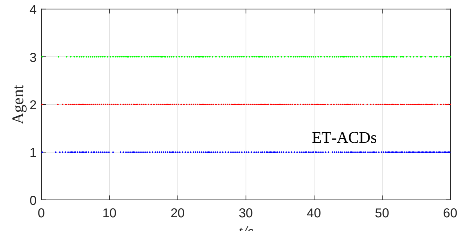

Fig. 14. The event-triggered instants under ET-ACDs

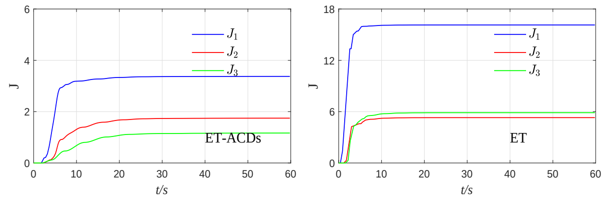

Fig. 15. Cumulative costs under ET-ACDs and ET strategies

Remark 11 Since many fields are inevitably influenced by stochastic factors in the real world, the autonomous cooperative control of many unmanned systems is more suitable to be modeled as the mean-square consensus problem of the SMASs. At the same time, due to the complexity and diversity of tasks undertaken by the multiple one-link manipulators, higher requirements are put forward for the control strategy design. In this paper, the optimal mean-square consensus of SMASs with input constraints is studied and the corresponding approximate event-triggered optimal controller is designed via the identifier-critic algorithm, which makes the control system more efficient and energy saving.

6 Conclusion

In this paper, the observer-based ETOC for nonlinear SMASs with unknown internal states and input constraints has been studied, and the event-triggered stochastic HJB equation with input constraints has been presented for the first time. The observer-based identifier network has been established to recover the knowledge of unknown system dynamics. Then, the approximate event-triggered optimal controller has been designed via ACDs. It is worth mentioning that there has not been literature on the ETOC for nonlinear stochastic SMASs with unknown internal states and input constraints via the framework of ET-ACDs. Moreover, it has been proved that the Zeno behavior does not exist in the closed-loop system. Finally, two simulation examples have been presented to further verify the validity of the ETOC scheme.

Funding This study was funded by the National Science Foundation of China (Grant No. 61773152).

Data availability statements The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Cui, M., Xie, X., Wu, Z.: Dynamics modeling and tracking control of robot manipulators in random vibration environment. IEEE Trans. Autom Control. **58**(6), 1540–1545 (2013)
2. Zhu, J., Lu, J., Yu, X.: Flocking of multi-agent nonholonomic systems with proximity graphs. IEEE Trans. Circuits Syst I, Reg. Papers. **60**(1), 199–210 (2013)
3. Dong, X., Li, Y., Lu, C., Hu, G., Li, Q., Ren, Z.: Time-varying formation tracking for UAV swarm systems with switching directed topologies. IEEE Trans. Neural Netw. Learn. Syst. **30**(12), 3674–3685 (2019)
4. Xu, S., Cao, J., Liu, Q., Rutkowski, L.: Optimal control on finite-time consensus of the leader-following stochastic multiagent system with heuristic method. IEEE Trans. Syst. Man Cybern. Syst. **51**(6), 3617–3628 (2021)
5. Hu, A., Cao, J., Hu, M., Guo, L.: Consensus of a leader-following multi-agent system with negative weights and noises. IET Control Theory Appl. **8**(2), 114–119 (2014)
6. Tang, Y., Gao, H., Zhang, W., Kurths, J.: Leader-following consensus of a class of stochastic delayed multi-agent systems with partial mixed impulses. Automatica. **53**, 346–354 (2015)
7. Tan, X., Cao, J., Li, X., Alsaedi, A.: Leader-following mean square consensus of stochastic multi-agent systems with input delay via event-triggered control. IET Control Theory Appl. **12**(2), 299–309 (2018)
8. Ren, H., Deng, F.: Mean square consensus of leader-following multi-agent systems with measurement noises and time delays. ISA Trans. **71**, 76–83 (2017)
9. Wei, Q., Wang, X., Zhong, X., Wu, N.: Consensus control of leader-following multi-agent systems in directed topology with heterogeneous disturbances. IEEE/CAA J. Autom. Sin. **8**(2), 423–431 (2021)
10. Werbos, P.: Advanced forecasting methods for global crisis warning and models of intelligence. General Syst. Yearbook. **22**(6), 25–38 (1977)
11. Zhang, H., Zhang, J., Yang, G., Luo, Y.: Leader-based optimal coordination control for the consensus problem of multiagent differential games via fuzzy adaptive dynamic programming. IEEE Trans. Fuzzy Syst. **23**(1), 152–163 (2015)
12. Zhong, X., He, H.: GrHDPs solution for optimal consensus control of multiagent discrete-time systems. IEEE Trans. Syst. Man Cybern. Syst. **50**(7), 2362–2374 (2020)
13. Lewis, F.L., Vrabie, D.: Reinforcement learning and adaptive dynamic programming for feedback control. IEEE Circuits Syst. Mag. **9**(3), 32–50 (2009)
14. Wang, F., Zhang, H., Liu, D.: Adaptive dynamic programming: An introduction. IEEE Comput. Intell. Mag. **4**(2), 39–47 (2009)
15. Kiumarsi B., Vamvoudakis, K.G., Modares, H., Lewis, F.L.: Optimal and autonomous control using reinforcement learning: A survey. IEEE Trans. Neural Netw. Learn. Syst. **29**(6), 2042–2062 (2018)
16. Liu, D., Xue, S., Zhao, B., Luo, B., Wei, Q.: Adaptive dynamic programming for control: A survey and recent advances. IEEE Trans. Syst. Man Cybern. Syst. **51**(1), 142–160 (2021)
17. Heemels, W., Johansson, K., Tabuada, P.: An introduction to event-triggered and self-triggered control. In: 2012 IEEE 51st IEEE Conference on Decision and Control (CDC). pp. 3270–3285 (2012)
18. Vamvoudakis, K.G.: Event-triggered optimal adaptive control algorithm for continuous-time nonlinear systems. IEEE/CAA J. Autom. Sin. **1**(3), pp. 282–293 (2014)
19. Zhao, W., Zhang, H.: Distributed optimal coordination control for nonlinear multi-agent systems using event-triggered adaptive dynamic programming method. ISA Trans. **91**, 184–195 (2019)
20. Sun, J., Long, T.: Event-triggered distributed zero-sum differential game for nonlinear multi-agent systems using adaptive dynamic programming. ISA Trans. **110**, 39–52 (2021)
21. Zhang, G., Zhu, Q.: Event-triggered optimal control for nonlinear stochastic systems via adaptive dynamic programming. Nonlinear Dyn. **105**(1), 387–401 (2021)
22. Dong, L., Zhong, X., Sun, C., He, H.: Event-triggered adaptive dynamic programming for continuous-time systems with control constraints. IEEE Trans. Neural Netw. Learn. Syst. **28**(8), 1941–1952, (2017)
23. Wang, D., Mu, C., Yang, X., Liu, D.: Event-based constrained robust control of affine systems incorporating an adaptive critic mechanism. IEEE Trans. Syst. Man Cybern. Syst. **47**(7), 1602–1612 (2017)
24. Yang, X., He, H.: Event-triggered robust stabilization of nonlinear input-constrained systems using single network adaptive critic designs. IEEE Trans. Syst. Man Cybern. Syst. **50**(9), 3145–3157, (2020)
25. Shi, Z., Zhou, C.: Distributed optimal consensus control for nonlinear multi-agent systems with input saturation based on event-triggered adaptive dynamic programming method. Int. J. Control. **95**(2), 282–294 (2022)
26. Zhu, Y., Zhao, D., He, H., Ji, J.: Event-triggered optimal control for partially unknown constrained-input systems via adaptive dynamic programming. IEEE Trans. Ind. Electron. **64**(5), 4101–4109 (2017)
27. Zhang, H., Zhang, K., Xiao, G., Jiang, H.: Robust optimal control scheme for unknown constrained-input nonlinear systems via a plug-n-play event-sampled critic-only algorithm. IEEE Trans. Syst. Man Cybern. Syst. **50**(9), 3169–3180 (2020)
28. Huo, X., Karimi, H.R., Zhao, X., Wang, B., Zong, G.: Adaptive-critic design for decentralized event-triggered control of constrained nonlinear interconnected systems within an identifier-critic framework. IEEE Trans. Cybern. (2021) https://doi: 10.1109/TCYB.2020.3037321
29. Zhang, H., Park, J.H. , Yue, D., Dou C.: Data-driven optimal event-triggered consensus control for unknown nonlinear multiagent systems with control constraints. Int. J. Robust Nonlinear Control. **29**(14), 4828–4844 (2019)
30. Ding, D., Wang, Z., and Han, Q., Neural-network-based consensus control for multiagent systems with input constraints: The event-triggered case. IEEE Trans. Cybern. **50**(8), 3719–3730 (2020)

31. Mao, X.: Stochastic versions of the lasalle theorem. J. Dyn. Differ. Equ. **153**(1), 175–195 (1999)
32. Tong, S., Li, Y., Li, Y., and Liu, Y.: Observer-based adaptive fuzzy backstepping control for a class of stochastic nonlinear strict-feedback systems. IEEE Trans. Syst. Man Cybern. Syst. **41**(6), 1693–1704 (2011)
33. Deng, H., Krstic, M., Williams, R.: Stabilization of stochastic nonlinear systems driven by noise of unknown covariance. IEEE Trans. Autom Control. **46**(8), 1237–1253 (2001)
34. Liu, C., Liu, L., Cao, J., Abdel-Aty, M.: Intermittent event-triggered optimal leader-following consensus for nonlinear multi-agent systems via actor-critic algorithm. IEEE Trans. Neural Netw. Learn. Syst. (2021) https://doi.org/10.1109/TNNLS.2021.3122458
35. Li, Y., Zhang, J., Liu, W., Tong, S.: Observer-based adaptive optimized control for stochastic nonlinear systems with input and state constraints. IEEE Trans. Neural Netw. Learn. Syst. (2021) https://doi: 10.1109/TNNLS.2021.3087796
36. Yong, J., Xun, Y.: Stochastic Controls: Hamiltonian Systems and HJB Equations. Springer. (1999)
37. Fabbri, G., Gozzi, F., Swiech, A.: Stochastic Optimal Control in Infinite Dimensions - Dynamic Programming and HJB Equations. HAL, Post-Print hal-01505767, (2017)
38. Liu, D., Wang, D., Wang, F., Li, H., Yang, X.: Neural-network-based online HJB solution for optimal robust guaranteed cost control of continuous-time uncertain nonlinear systems. IEEE Trans. Cybern. **44**(12), 2834–2847 (2014)
39. Jiao, Q., Modares, H., Xu, S., Lewis, F.L., Vamvoudakis, K.G.: Multi-agent zero-sum differential graphical games for disturbance rejection in distributed control. Automatica. **69**, 24–34 (2016)
40. Su, H., Zhang, H., Gao, D., Luo Y.: Adaptive dynamics programming for H_∞ control of continuous-time unknown nonlinear systems via generalized fuzzy hyperbolic models. IEEE Trans. Syst. Man Cybern. Syst. **50**(11), 3996–4008 (2020)
41. Sahoo, A., Xu, H., Jagannathan, S.: Neural network-based event-triggered state feedback control of nonlinear continuous-time systems. IEEE Trans. Neural Netw. Learn. Syst. **27**(3), 497–509 (2016)
42. Wen, G., Xu, L., Li B.: Optimized backstepping tracking control using reinforcement learning for a class of stochastic nonlinear strict-feedback systems. IEEE Trans. Neural Netw. Learn. Syst. (2021) https://doi: 10.1109/TNNLS.2021.3105176
43. Ren, C., Chen, L., Chen, C.L.P.: Adaptive fuzzy leader-following consensus control for stochastic multiagent systems with heterogeneous nonlinear dynamics. IEEE Trans. Fuzzy Syst. **25**(1), 181–190 (2017)