Event-Driven-Modular Adaptive Backstepping Optimal Control for Strict-Feedback Systems Through Zero-Sum Differential Games

YUEHUI JI1,2, HAILIANG ZHOU3, AND BING BA1,2
1School of Electrical and Electronic Engineering, Tianjin University of Technology, Tianjin 300384, China
2Tianjin Key Laboratory for Control Theory and Applications in Complicated Systems, Tianjin University of Technology, Tianjin 300384, China
3Tianjin Institute of Metrological Supervision and Testing (TIMST), Tianjin 300192, China

Corresponding author: Yuehui Ji (jiyuehuitju@126.com)

This work was supported in part by the Research Project of Tianjin Municipal Education Commission under Grant 2017KJ249, and in part by the National Natural Science Foundation of China under Grant 61673294.

ABSTRACT This paper addresses the event-driven-modular optimal tracking control problem for nonlinear strict-feedback systems with external disturbances. Through the backstepping feedforward control, the optimal tracking problem is transformed into an equivalent optimal regulation problem of affine tracking error system. Subsequently, adaptive dynamic programming technique is introduced to generate the optimal feedback controller, and solve the optimization problem of two-player zero-sum differential game. A single critic neural network is constructed to evaluate the associated cost function online, where the novel weight updating law is derived based on the gradient-descent technique. The resulting event-triggered closed-loop system, modeled as an impulsive system, is proved to be asymptotically stable by Lyapunov theory. Finally, the reliability and effectiveness of the theoretical results is validated by numerical simulation examples.

INDEX TERMS Adaptive backstepping optimal, event triggering control, zero-sum differential game, adaptive dynamic programming (ADP), neural network (NN).

I. INTRODUCTION

During the past decades, research on nonlinear strict-feedback systems have drawn considerable attention, i.e., hypersonic flight vehicle [1], inverted pendulum system [2], quadrotor [3], helicopter [4] ship autopilot [5], and robot manipulator [6], as the nonlinear general system satisfying certain geometric conditions can be transformed into strict-feedback form by the diffeomorphism theory. The celebrated backstepping recursive technology has been studied to accomplish the tracking control issue for strict-feedback systems [7]–[12]. Considering high-order nonlinear multiagent systems in semi-strict-feedback form, the neural network state observer is constructed for each follower and an adaptive consensus tracking control is studied via backstepping techniques [7]. A state-feedback control is constructed by backstepping, rational-exponent Lyapunov functions and Bernoulli inequality for nonlinear strict-feedback system with reduced design complexity [8]. The adaptive backstepping control is investigated for uncertain systems subjected to input delay and disturbances. To compensate the input delay, pade approximation method is introduced to construct an auxiliary system [9]. Considering the magnetic levitation systems, nonlinear integral backstepping controllers are designed to generate magnetic flux to levitate a ferromagnetic object in the air [10]. For a special class of nonlinear strict-feedback systems performing the same operation, the echo state network and backstepping are combined to accomplish a iterative learning control scheme [11]. The backstepping control technique is present for a rotary inverted pendulum, only the system structure and state measurements available [12]. Despite of efforts, system optimality is not covered in the aforementioned research.

The optimal control can greatly promote social development and national economic construction; hence system optimality is a higher priority than the stability in practical engineering and some important results have been achieved recently. In the traditional optimal control, solving the Hamilton-Jacobi-Isaacs (HJI) equation is a intractable issue as a general closed-form analytical method has not yet been developed [13], [14]. To solve the HJI equation effectively, an adaptive optimal reinforcement learning control is...
investigated for discrete-time systems based on backstepping and minimal learning parameter technique [15]. The ADP-based $H_{\infty}$ tracking control is studied for discrete time-delay systems, deriving a data type Bellman equation [16]. The model-free state feedback tracking control is accomplished via approximate dynamic programming with the value iteration algorithm [17]. A data-driven trajectory tracking control is addressed for nonlinear discrete-time systems in presence of unknown dynamics [18]. In view of the strict feedback system, adaptive dynamic programming (ADP) is proposed to design the optimal control and minimize some predefined performance cost function. An optimal tracking control scheme, which is composed of backstepping-generated adaptive feedforward control and dynamic programming-based optimal feedback control, is proposed for continuous-time strict feedback systems [19]. An adaptive fuzzy decentralized optimal control is investigated for nonlinear large-scale systems subjected to the unknown nonlinear functions and unmeasured states [20]. For multi-agent systems in strict-feedback form with a fixed directed graph, the command-filtered backstepping and adaptive dynamic programming technique are introduced to investigate the distributed fuzzy optimal tracking control [21]. However, the ADP control schemes [19]–[21] mentioned above are restricted to the assumptions that the system is not affected by external disturbances.

As a main branch of game theory, differential game theory introduces differential equations to address the issue of dynamic conflicts, competition or cooperation for multiple objects [22], [23]. In the zero-sum (ZS) differential game, the control policies are constructed to minimize a predefined cost function corresponding to the worst-case disturbance, and to obtain the saddle-point equilibrium solution. Exactly, the problem of attenuating the effect of external disturbances for nonlinear strict-feedback systems is a two-player zero-sum differential game problem, where the control input is designed to minimize the performance cost function, and the disturbance policy is generated to maximize the performance cost function. The thought is succeeded for intercepting high-speed manoeuvring targets of interceptor missiles with output and input constraints [24], and stimulate the choice of optimal ADP zero-sum control strategy for the tracking error dynamics in nonlinear strict-feedback systems.

To reduce the computational burden and/or interactive information in plant-controller communication networks, developing an event-triggering control (ETC) is crucial scheme to decrease the power consumption of the actuators batteries and slowed down the actuators abrasion simultaneously [25]–[27]. In event-sampling control systems, the control input is updated only at a sequence of time instants, which are determined by the significant event-triggered condition. In general, the aperiodic updates or transmissions depend on the current state of the plant, which is more efficient than the periodic time-triggered execution. Referring to backstepping method in nonlinear strict-feedback systems, the adaptive event-triggered control is concerned by exploiting the event-sampled neural network and backstepping method [28]. An adaptive backstepping control is proposed for parametric strict-feedback systems, where the parameter estimator and controller are updated at the event-triggered instants [29]. The adaptive fuzzy control based on an event-triggered mechanism is investigated for nonlinear strict-feedback systems in the presence of virtual control coefficients and actuator failures [30]. For event-triggered optimal control, the optimization problem of zero-sum differential game is solved in the framework of event triggering and adaptive dynamic programming algorithm [31]. The zero-sum game problem is addressed for partially unknown drift system, and an event-triggered adaptive dynamic programming method is developed [32]. To the best of our knowledge, there are few reports on event-triggered control of adaptive backstepping optimal control. In strict-feedback systems, the topic of optimization performance and energy-saving property for the control system is interesting and challenging, which is the motivation of the paper.

Inspired by the existing research, an adaptive event-sampling algorithm is devised and a two-player ZS differential game is formulated for strict-feedback systems in the presence of external disturbances. The whole controller in the framework of ET mechanism consists of a feedforward tracking controller by backstepping and an optimal feedback control by ADP algorithm. The main novelty of the paper can be extracted as: (I) A strict-feedback system in presence of external disturbances is converted into an equivalent two-player zero-sum differential game by backstepping. For the resulting affine tracking error system, the optimal feedback control is fulfilled by ADP technique. Compared to the research in [28]–[30], the proposed adaptive control can ensure the stability as well as the the optimality of the closed-loop system. (II) An event-triggered condition is specified to drive signal sampling and controller execution for two-player ZS games, which constitutes both the event-triggered control and time-triggered disturbance. In contrast to work [24], event-triggered control can achieve less communication resources and/or computational burden effectively.

The problem formulation is presented in Section II. The feedforward tracking controller is proposed in Section III. Section IV formulates the optimal feedback controller, including the implementation of single adaptive critic neural network and event-triggered generator. Rigorous mathematical stability analysis of closed-loop system is accomplished in Section V. Two simulation results are provided to illustrate the effectiveness of the proposed control scheme in Section VI. Finally, the conclusion in Section VII summarizes the paper.

II. PROBLEM STATEMENT

The nonlinear strict-feedback system dynamic in presence of external disturbance is described as:

\[
\begin{align*}
\dot{x}_i(t) &= f_i(\tilde{x}_i(t)) + g_i x_{i+1}(t) + d_i(t), \quad i = 1, 2 \ldots n - 1 \\
\dot{x}_n(t) &= f_n(x(t)) + g_n u(t) + d_n(t) \\
y(t) &= x_1(t)
\end{align*}
\]
where \( x(t) = [x_1(t), \ldots, x_i(t), \ldots, x_n(t)]^T \in \mathbb{R}^n \), \( u(t) \in \mathbb{R}, y(t) \in \mathbb{R} \) denote system state vector, control input and available system output, respectively; \( \hat{x}_i(t) = [\hat{x}_1(t), \ldots, \hat{x}_i(t)] \in \mathbb{R}^i \) is partial state vector; \( f_i(\hat{x}_i(t)) \in \mathbb{R}^i \) is the continuous function and \( f_i(0) = 0 \) is known constant with the common assumption that \( g_i \geq 0; d_i(t) \in \mathbb{R} \) is an unknown and bounded disturbance. The aforementioned variables and vectors are related to time \( t \). For concise, time \( t \) will be omitted in the later description.

**Assumption 1:** The reference signal \( x_{id}(t) \in \mathbb{R} \) and its first-order derivative \( \dot{x}_{id}(t) \in \mathbb{R} \) are available, continuous and bounded.

The control objective is to construct an optimal control in event-triggered mechanism for strict-feedback system (1), such that the system output is driven to track the desired reference signal \( x_{id}(t) \) in an optimal manner and all the signals in closed-loop system are uniformly ultimately bounded (UUB).

### III. FEEDFORWARD TRACKING CONTROL

The celebrated backstepping technology is introduced to design a feedforward controller, and transform the tracking problem for nonlinear strict-feedback system (1) into an optimal regulation problem for system state tracking errors. The feedforward controller is designed based on the following coordinate transformation:

\[
\begin{align*}
z_1 &= x_1 - x_{id}; \\
z_i &= x_i - x_{id}; & i = 2, \ldots, n
\end{align*}
\]

with \( z_i, i = 1, \ldots, n \) is the state tracking error and \( x_{id}, i = 2, \ldots, n \) is the stabilized control. The control law \( x_{id} \) consists of feedforward virtual control \( u_{id}^{b-1} \) by backstepping technology and optimal feedback control \( u_{id}^{f-1} \) by ADP:

\[
x_{id} = u_{id}^{b-1} + u_{id}^{f-1}
\]

To implement the event-triggered control, a monotonical increasing time sequence \( \{t_s\}_{s=0}^{\infty} \) is defined as the event-triggered sampling instant satisfying \( t_{s+1} > t_s, t_s \in \mathbb{R}_0^+ \), \( \forall s \in Z^+ \), and \( t_0 = 0 \) is the initial sampling instant. At the instants \( t = t_s \), the event generator(EG) is triggered and the last held states are updated with the current system states. Accordingly, the control input is updated at the trigger instants \( t = t_s \), and held by a zero-order holder (ZOH) until the next trigger instant comes. Therefore, the piecewise continuity of the control input can be maintained through the “EG-ZOH” mechanism. The last held state is denoted as:

\[
\hat{x}_i(t) = x_i(t_s), \quad t = t_s, \quad \forall s \in Z^+, \quad i = 1, 2, \ldots, n
\]

The event-triggered error between the current state and the last held state is:

\[
e_i(t) = x_i(t) - \hat{x}_i(t), \quad \forall t \in [t_s, t_{s+1}], \quad i = 1, 2, \ldots, n
\]

Defined event-referred tracking error as:

\[
\begin{align*}
\hat{z}_1 &= \hat{x}_1 - x_{id}; \\
\hat{z}_i &= \hat{x}_i - x_{id}; & i = 2, \ldots, n
\end{align*}
\]

The event-triggered feedforward controller laws, predicated on the sampled state \( \hat{x}_i(t) \) instead of the real state \( x_i(t) \), are constructed as:

\[
\begin{align*}
u_{id}^f &= g_i^{-1}(-k_1\hat{z}_1 - f_i(x_{id}) - \hat{x}_{id}) \\
u_{id}^b &= g_i^{-1}(-k_i\hat{z}_i - f_i(\hat{x}_i)), & 2 \leq i < n \\
u_{id}^b &= g_n^{-1}(-k_n\hat{z}_n - f_n(\hat{x}_{nd}))
\end{align*}
\]

where \( k_i \) is controller gain.

**Step 1:** The dynamic of tracking error \( z_1 \) is:

\[
\begin{align*}
\dot{z}_1 &= \dot{x}_1 - \dot{x}_{id} = f_1(x_1) + g_1x_2 + d_1 - \dot{x}_{id} \\
&= f_1(x_1) + g_1(z_2 + u_1^b + u_1^f) + d_1 - \dot{x}_{id}
\end{align*}
\]

Considering the feedforward virtual control (7), we can deduce that:

\[
\begin{align*}
\dot{z}_1 &= -k_1\dot{x}_1 + g_1z_2 + f_1(x_1) - f_1(x_{id}) + g_1u_1^a + d_1 \\
&= -k_1\dot{z}_1 + k_1e_1 + g_1z_2 + H_1(z_1) + g_1u_1^a + d_1 + \dot{x}_{id}
\end{align*}
\]

where \( H_1(z_1) \triangleq f_1(x_1) - f_1(x_{id}) \).

**Step 2:** The dynamic of \( z_i \) is:

\[
\begin{align*}
\dot{z}_i &= \dot{x}_i - \dot{x}_{id} = f_i(x_i) + g_ix_{i+1} + d_i - \dot{x}_{id} \\
&= f_i(x_i) + g_i(z_{i+1} + u_i^b + u_i^f) + d_i - \dot{x}_{id}
\end{align*}
\]

Substituting the feedforward virtual control (7) into Eqns. (10), yields that:

\[
\begin{align*}
\dot{z}_i &= -k_i\dot{x}_i + g_iz_{i+1} + f_i(x_i) - f_i(x_{id}) + g_iu_i^a + d_i - \dot{x}_{id} \\
&= -k_i\dot{z}_i + k_ie_i + g_iz_{i+1} + H_i(z_i) + g_iu_i^a + d_i + \dot{x}_{id}
\end{align*}
\]

where \( H_i(z_i) \triangleq f_i(x_i) - f_i(x_{id}) \) and \( \hat{z}_i = [z_1, z_2, \ldots, z_i]^T \).

**Step n:** The dynamic of \( z_n \) is:

\[
\dot{z}_n = \dot{x}_n - \dot{x}_{nd} = f_n(x) + g_nu + d_n - \dot{x}_{nd}
\]

Considering the feedforward control (7), yields that:

\[
\begin{align*}
\dot{z}_n &= f_n(x) + g_nu^b + g_nu^a + d_n - \dot{x}_{nd} \\
&= f_n(x) - k_n\dot{z}_n + f_n(x_{nd}) + g_nu^a + d_n - \dot{x}_{nd} \\
&= -k_n\dot{z}_n + k_ne_n + H_n(z_n) + g_nu^a + d_n - \dot{x}_{nd}
\end{align*}
\]

where \( H_n(z_n) \triangleq f_n(x_n) - f_n(x_{nd}) \).

During the flow interval \( t \in [t_s, t_{s+1}] \), the control \( x_{id} \) are held by ZOH, and \( \dot{x}_{id} = 0 \). The dynamic of tracking error \( \hat{z}_i \) is:

\[
\begin{align*}
\dot{z}_1 &= -k_1z_1 + k_1e_1 + g_1z_2 + H_1(z_1) + g_1u_1^a + d_1 \\
\dot{z}_i &= -k_i\hat{z}_i + k_i\hat{z}_i + g_i\hat{z}_{i+1} + H_i(z_i) + g_iu_i^a + d_i \\
\dot{z}_n &= -k_n\hat{z}_n + k_n\hat{z}_n + H_n(z_n) + g_nu_n^a + d_n
\end{align*}
\]

The Lyapunov function candidate \( V_z \) is defined as:

\[
V_z = \sum_{i=1}^{n} \hat{z}_i^2 / 2
\]

The first-time derivative of \( V_z \) along the trajectory (14) can be expressed as:

\[
\dot{V_z} = \sum_{i=1}^{n} z_i\dot{z}_i = \sum_{i=1}^{n} \left( -k_i\hat{z}_i^2 + k_i\hat{z}_i e_i + g_i\hat{z}_i\hat{z}_{i+1} + H_i(z_i) + z_i H_i(z_i) + z_i g_i u_i^a + z_i d_i \right)
\]

where \( z_{i+1} = 0 \).
By Young’s inequality, we can obtain that:
\[ z_{i-1} \leq \frac{1}{2} z_{i}^2 + \frac{1}{2} z_{i+1} \]  
(17)

Substituting Eqs.(17) into Eqs.(16), we have:
\[ \dot{V}_Z = \sum_{i=1}^{n} \left( \left( k_i - \frac{1}{2} g_i - \frac{1}{2} g_{i-1} \right) z_i^2 + k_i z_{i+1} \right) + Z^T H(Z) + Z^T G_a U^a + Z^T G_d D \]  
(18)

with \( g_0 = 0, Z = [z_1, z_2, \ldots, z_n]^T; G_a = \text{diag} \{g_1, \ldots, g_n\} \)

The control problem for nonlinear strict-feedback system with external disturbance (1) has been transformed into the equivalent regulation problem for tracking error (19). Then the ADP-based differential game scheme will be introduced to stabilize the affine system (19) and achieve the optimal control performance.

IV. OPTIMAL FEEDBACK CONTROL

The affine system dynamic of tracking error \( Z \) is described as:
\[ \dot{Z} = H(Z) + G_a U^a + G_d D \]  
(19)

where the optimal feedback control vector \( U^a \) and disturbance input vector \( D \) are considered as two control inputs.

The non-quadratic performance cost function is defined as:
\[ J = \int_{t=0}^{\infty} r(Z(\tau), U^a(\tau), D(\tau))d\tau \]  
(20)

Therefore, the control problem can be viewed a two-player zero-sum optimal control problem, i.e., the feedback control policy \( U^a \) is regarded as player 1 and the optimal control input is sought to minimize the performance cost function (20); while the disturbance policy \( D \) is considered as player 2 and the worst-case disturbance input is found to maximize the performance cost function (20). Herein, the utility function \( r(Z(\tau), U^a(\tau), D(\tau)) \) is defined as:
\[ r(Z(\tau), U^a(\tau), D(\tau)) = Z^T Q(Z) + (U^a)^T (H(Z) + G_a U^a + G_d D) + Z^T Q(Z) + (U^a)^T \Sigma U^a - \eta^2 D^T D \]  
(21)

where \( Q, \Sigma \) are positive definite symmetric matrices and \( \eta > 0 \) is the disturbance attenuation constant.

Definition 1: A set of control pairs \( \{U^a, D^*\} \) is the saddle-point equilibrium for the two-player zero-sum game, if the following inequalities hold:
\[ J(U^a, D^*) \geq J(U^a, D) \geq J(U^a, D^*) \]  
(22)

The corresponding Hamilton function is defined as:
\[ h(Z, U^a, D, J) = (\nabla J(Z))^T (H(Z) + G_a U^a + G_d D) + Z^T Q(Z) + (U^a)^T \Sigma U^a - \eta^2 D^T D \]  
(23)

According to Bellman’s principle of optimality, the optimal performance cost function \( J^*(Z) \) satisfies that:
\[ \min_{U^a} \max_{D} h(Z, U^a, D, J^*) = \max_{D} \min_{U^a} h(Z, U^a, D, J^*) = 0 \]  

The Hamilton-Jacobi-Isaacs (HJI) equation and the stationary conditions hold:
\[ h(Z, (U^a)^*, (D^*)^*, J^*) = (\nabla J^*(Z))^T (H(Z) + G_a U^a + G_d D) + Z^T Q(Z) + (U^a)^T \Sigma U^a - \eta^2 D^T D = 0 \]  
(24)

Then the optimal-control/worst-case disturbance pair \( \{(U^a)^*, (D^*)^*\} \) is chosen as:
\[ U^a = \frac{1}{2} \Sigma^{-1} G_a^T \nabla J^*(Z) \]  
(25)

Then the HJI equation (24) under the control input (25) can be rewritten as:
\[ 0 = Z^T Q(Z) + (\nabla J^*(Z))^T (H(Z) + G_a U^a + G_d D) \]  
(26)

Lemma 1: Consider the tracking error system (19), the associated performance cost function (20), and the input pair (25). Assume that there exists a continuous differentiable Lyapunov function \( J_s(Z) \) such that \( J_s(Z) = (\nabla J_s(Z))^T Z = (\nabla J_s(Z))^T \tilde{Z} = (\nabla J_s(Z))^T (H(Z) + G_a U^a + G_d D) \leq 0 \), where \( \nabla J_s(Z) \) is gradient of \( J_s(Z) \) with respect to \( Z \). Moreover, let \( \lambda(Z) \) be a positive definite function, i.e. \( \forall Z \neq 0, \lambda(Z) > 0 \), and \( \lambda(Z) \rightarrow \infty \) as \( Z \rightarrow 0 \). Furthermore, let \( \lambda(Z) \) satisfies that \( \lim_{Z \rightarrow \infty} \lambda(Z) = \infty \) as well as:
\[ (\nabla J(Z))^T \lambda(Z) \nabla J_s(Z) = Z^T Q(Z) + (U^a)^T \Sigma U^a - \eta^2 D^T D \]  
(27)

Then, the following inequation holds:
\[ (\nabla J_s(Z))^T (H(Z) + G_a U^a + G_d D) < -(\nabla J_s(Z))^T \lambda(Z) \nabla J_s(Z) \]  
(28)

Assumption 2: Assume that the optimal closed-loop system is bounded by the function of system states, such that \( ||H(Z) + G_a U^a + G_d D|| \leq c \sqrt{\|Z\|^2} \), with \( c > 0 \).

The event-triggered optimal feedback control input is proposed as:
\[ U^a(Z(t)) = \begin{cases} -\frac{1}{2} \Sigma^{-1} G_a^T \nabla J^*(Z(t)) = -\frac{1}{2} \Sigma^{-1} G_a^T \nabla J^*(\tilde{Z}), & t = t_s \\ (U^a)^*(Z(t)) = (U^a)^*(\tilde{Z}), & t \in (t_s, t_{s+1}) \end{cases} \]  
(29)
where \( e = [e_1, e_2, \ldots, e_d]^T \) is event-triggered error vector, and \( \hat{Z} \) is last held tracking error vector defined as:

\[
\hat{Z} = Z - e = \begin{cases} Z, & t = t_s \\
Z(t_s), & t \in (t_s, t_{s+1}) \end{cases}
\]

Therefore, the piecewise continuity of the control input can be maintained through the “EG-ZOH” mechanism. As an external signal to suppress, the time-driven disturbance law, related to \( \hat{Z}(t) \), is defined as:

\[
D^*(Z(t)) = \frac{1}{2\eta^2} G_d^2 \nabla J^*(Z(t))
\]

(30)

Subsequently, the event-triggered system dynamic for the tracking error is described as:

\[
\dot{\hat{Z}} = H(Z) + G_u U^a(\hat{Z}) + G_d D^*(Z)
= H(Z) + G_u U^a(\hat{Z} - e) + G_d D^*(Z)
\]

(31)

A. ADAPTIVE CRITIC NEURAL NETWORK FRAMEWORK

The solution of HJI equation (26) is the primary problem to implement the feedback control/disturbance policy \( ([U^a]^*, D^*) \), which provide a saddle-point for the zero-sum differential games. However, to obtain the analytic solution of the HJI equation (26) is generally difficult even impossible, due to the nonlinear features in tracking error dynamic (19). Subsequently, the reinforcement learning mechanism is utilized to construct the critic network and solve the HJI equation (26) approximately in the zero-sum optimal control problem. A three-layer feedforward neural network is introduced to reconstruct the performance cost function \( J^*(Z) \) on a compact set \( \Omega \):

\[
J^*(Z) = W_c^T \psi_c(Z) + e_c(Z)
\]

(32)

where \( W_c \in \mathbb{R}^{N_c} \) is the ideal bounded weight vector, \( \psi_c(Z) : \mathbb{R}^m \rightarrow \mathbb{R}^{N_c} \) is the activation function, \( N_c \) is the neurons quantity in the hidden layer, \( e_c(Z) \in \mathbb{R} \) is the finite approximation error.

The gradient vector of the performance cost function with respect to \( Z \) is:

\[
\nabla J^*(Z) = \nabla \psi_c^T(Z) W_c + \nabla e_c(Z)
\]

(33)

The event-triggered optimal control/time-triggered worst-case disturbance pair can be deduced as:

\[
(U^a)^*(\hat{Z}) = -\frac{1}{2} \xi^{-1} G_u^T \left( \nabla \psi_c^T(\hat{Z}) W_c + \nabla e_c(\hat{Z}) \right)
\]

\[
D^*(Z) = \frac{1}{2\eta^2} G_d^2 \left( \nabla \psi_c^T(Z) W_c + \nabla e_c(Z) \right)
\]

(34)

The HJI equation (26) can be derived as:

\[
0 = W_c^T \nabla \psi_c(Z) H(Z) + Z^T Q Z + \varepsilon_H
- \frac{1}{2} W_c^T \nabla \psi_c(Z) \Gamma_a \nabla \psi_c^T(\hat{Z}) W_c
+ \frac{1}{4} W_c^T \nabla \psi_c(Z) \Gamma_a \nabla \psi_c^T(\hat{Z}) W_c
+ \frac{1}{2} W_c^T \nabla \psi_c(Z) \Gamma_d \nabla \psi_c^T(Z) W_c
\]

(35)

where \( \Gamma_a = G_u \xi^{-1} G_u^T, \Gamma_d = G_d G_d^T / \eta^2 \) and the residual error \( \varepsilon_H \) due to the function approximation error is:

\[
\varepsilon_H = \nabla \varepsilon_H^T(\hat{Z}) (H(Z) + G_u U^a(\hat{Z}) + G_d D^*)
+ \frac{1}{2} W_c^T \nabla \psi_c(Z) \Gamma_a \nabla \psi_c^T(\hat{Z}) \Gamma_a \nabla \psi_c(Z)
+ \frac{1}{4} W_c^T \nabla \psi_c(Z) \Gamma_a \nabla \psi_c^T(\hat{Z}) \Gamma_d \nabla \psi_c(Z)
+ \frac{1}{2} W_c^T \nabla \psi_c(Z) \Gamma_d \nabla \psi_c^T(\hat{Z}) \Gamma_a \nabla \psi_c(Z)
+ \frac{1}{4} W_c^T \nabla \psi_c(Z) \Gamma_d \nabla \psi_c^T(Z) W_c
+ \frac{1}{4} \nabla \varepsilon_H^T(\hat{Z}) \Gamma_d \nabla \psi_c(Z)
\]

(36)

As the ideal weight vector is unknown, the actual critic network output is built as:

\[
\hat{j}(Z) = \hat{W}_c \psi_c(Z)
\]

(37)

where \( \hat{W}_c \) denotes the estimated weight vector.

Subsequently, the gradient vector can be expressed as:

\[
\nabla \hat{j}(Z) = \nabla \psi_c^T(\hat{Z}) \hat{W}_c
\]

(38)

The approximated optimal control and worst-case disturbance inputs are implemented as:

\[
\hat{U}^a(\hat{Z}) = \frac{1}{2} \xi^{-1} G_u^T \left( \nabla \psi_c^T(\hat{Z}) \hat{W}_c(t_s) \right)
\]

\[
\hat{D}(Z) = \frac{1}{2\eta^2} G_d^2 \left( \nabla \psi_c^T(Z) \hat{W}_c(t) \right)
\]

(39)

Define the estimation error of the weight vector in critic network as:

\[
\tilde{W}_c = W_c - \hat{W}_c
\]

(40)

Then the approximated Hamilton-Jacobi-Isaacs function yields that:

\[
\hat{h}(Z, \hat{Z}, \tilde{W}_c) = \hat{W}_c^T \nabla \psi_c(Z) \left( H(Z) + G_u \hat{U}^a(\hat{Z}) + G_d \hat{D}(Z) \right)
+ Z^T Q Z + (\hat{U}^a(\hat{Z}))^T \xi \hat{U}^a(\hat{Z}) - \eta^2 \hat{D}(Z) \dot{\hat{D}}(Z)
\]

\[
= \hat{W}_c^T \nabla \psi_c(Z) H(Z) + Z^T Q Z - \frac{1}{2} \hat{W}_c^T \nabla \psi_c(Z)
\times \Gamma_a \nabla \psi_c^T(\hat{Z}) \hat{W}_c(t_s)
+ \frac{1}{4} \hat{W}_c^T \nabla \psi_c(Z) \Gamma_a \nabla \psi_c^T(\hat{Z}) \hat{W}_c(t_s)
+ \frac{1}{4} \hat{W}_c^T \nabla \psi_c(Z) \Gamma_d \nabla \psi_c^T(Z) \hat{W}_c(t)
\]

\[
\triangleq \xi_H
\]

(41)

which yields that:

\[
\frac{\partial \xi_H}{\partial \hat{W}_c} = \nabla \psi_c(Z) \left( H(Z) + G_u \hat{U}^a(\hat{Z}) + G_d \hat{D}(Z) \right)
= \nabla \psi_c(Z) H(Z) - \frac{1}{2} \nabla \psi_c(Z) \Gamma_a \left( \nabla \psi_c^T(\hat{Z}) \hat{W}_c(t_s) \right)
+ \frac{1}{2} \nabla \psi_c(Z) \Gamma_d \left( \nabla \psi_c^T(Z) \hat{W}_c(t) \right) \triangleq \nu
\]

(42)
Considering (36), we have:

\[
Z^T Q Z = -W_c^T \nabla \phi_c(Z) H(Z) - \varepsilon_H + \frac{1}{2} W_c^T \nabla \phi_c(Z) \Gamma_u \nabla \phi_c^T(\hat{Z}) W_c + \frac{1}{4} W_c^T \nabla \phi_c(\hat{Z}) \Gamma_u \nabla \phi_c^T(\hat{Z}) W_c - \frac{1}{4} W_c^T \nabla \phi_c(\hat{Z}) \Gamma_d \nabla \phi_c^T(\hat{Z}) W_c
\]

and the approximated HJJ function satisfies:

\[
\xi_H = \frac{1}{2} W_c^T \nabla \phi_c(Z) H(Z) - \varepsilon_H + \frac{1}{2} W_c^T \nabla \phi_c(Z) \Gamma_u \nabla \phi_c^T(\hat{Z}) W_c - \frac{1}{2} W_c^T \nabla \phi_c(\hat{Z}) \Gamma_u \nabla \phi_c^T(\hat{Z}) W_c + \frac{1}{2} W_c^T \nabla \phi_c(\hat{Z}) \Gamma_d \nabla \phi_c^T(\hat{Z}) W_c - \frac{1}{4} W_c^T \nabla \phi_c(\hat{Z}) \Gamma_d \nabla \phi_c^T(\hat{Z}) W_c + \frac{1}{4} W_c^T \nabla \phi_c(\hat{Z}) \Gamma_d \nabla \phi_c^T(\hat{Z}) W_c - \frac{1}{4} W_c^T \nabla \phi_c(\hat{Z}) \Gamma_d \nabla \phi_c^T(\hat{Z}) W_c
\]

Note that the ideal value of approximated Hamilton function (42) is 0, when \( \hat{W}_c \to W_c \). The critic weight vector is trained to minimize the squared residual error \( E_c \):

\[
E_c = \frac{1}{2} \xi_H^T \xi_H
\]

By a modified normalized gradient-descent algorithm, the update law of the critic weight matrix is tuned as:

\[
\dot{\hat{W}}_c = -\frac{u_1}{u_2} \nabla \phi_c(Z) (\Gamma_u - \Gamma_d) \nabla J_c(Z) + \frac{1}{2} \nabla \phi_c(Z) (\Gamma_u + \Gamma_d) \nabla \phi_c^T(\hat{Z}) \hat{W}_c + \frac{1}{4} \nabla \phi_c(\hat{Z}) \Gamma_u \nabla \phi_c^T(\hat{Z}) \hat{W}_c - \frac{1}{4} \nabla \phi_c(\hat{Z}) \Gamma_d \nabla \phi_c^T(\hat{Z}) \hat{W}_c
\]

where \( u \) is the learning rate, \( u_1 \equiv u_1 / u_2 \equiv 1 + u^T u, u_2^2 = (1 + u^T u)^2 \) is the normalized term, \( \lambda_1, \lambda_2 \) is the tuning gains, the term \( J_c(Z) \) is defined in Lemma 1.

The weight error dynamics of the critic NN can be deduced as:

\[
\hat{W}_c = -\frac{u_1}{u_2} \nabla \phi_c(Z) H(Z) - \varepsilon_H + \frac{1}{2} W_c^T \nabla \phi_c(Z) \Gamma_u \nabla \phi_c^T(\hat{Z}) W_c + \frac{1}{4} W_c^T \nabla \phi_c(\hat{Z}) \Gamma_u \nabla \phi_c^T(\hat{Z}) W_c - \frac{1}{4} W_c^T \nabla \phi_c(\hat{Z}) \Gamma_d \nabla \phi_c^T(\hat{Z}) W_c - \frac{1}{4} W_c^T \nabla \phi_c(\hat{Z}) \Gamma_d \nabla \phi_c^T(\hat{Z}) W_c + \frac{1}{4} W_c^T \nabla \phi_c(\hat{Z}) \Gamma_d \nabla \phi_c^T(\hat{Z}) W_c - \frac{1}{4} W_c^T \nabla \phi_c(\hat{Z}) \Gamma_d \nabla \phi_c^T(\hat{Z}) W_c
\]

For the critic NN (32), the following assumption is provided.

**Assumption 3:**

1. In the critic NN, the activation function, the approximation error and the corresponding gradients with respect to \( Z \) are upper bounded, such that \( \| \phi_c(Z) \| \leq \hat{\phi}_c, \| \varepsilon_c(Z) \| \leq \hat{\varepsilon}_c, \| \nabla \phi_c(Z) \| \leq \hat{\phi}_c, \| \nabla \varepsilon_c(Z) \| \leq \hat{\varepsilon}_c, \| \phi_c(Z) \| \leq \hat{\phi}_c, \| \varepsilon_c(Z) \| \leq \hat{\varepsilon}_c \). where \( \hat{\phi}_c, \hat{\phi}_c, \hat{\varepsilon}_c, \hat{\varepsilon}_c, \hat{\phi}_c, \hat{\varepsilon}_c \) are the certain positive constants. The boundedness assumption also holds for \( \phi_c(Z), \varepsilon_c(Z), \phi_c(Z), \varepsilon_c(Z) \).

2. The ideal weight vector \( W_c \) and the estimated weight vector at the trigger instant \( \hat{W}_c(t_i) \) are upper bounded.

**B. EVENT GENERATOR DESIGN**

The traditional time-triggered control receives state-feedback information continuously and the control actions are updated periodically, resulting in unnecessary resource consumptions and communication costs. In this section, an event generator is implemented to reduce the tremendous state sampling, without compromising the system stability. The state feedback signals are transmitted, then the optimal controller is updated only when the event-triggered conditions are satisfied:

\[
\| e \|^2 \geq \max \left( 1 - \sigma \frac{k_{\min} - 1}{2\tau} - \frac{1}{2\lambda} k_{\max}, \frac{1}{2} \lambda k_{\max}, \rho \right)
\]

where \( k_{\min} = \min\{k_l - \frac{1}{\sigma^2}, 0\} \), \( k_{\max} = \max\{k_i\}, \sigma, \lambda, \) are proper parameters satisfying that \( 0 < \sigma < 1, \tau > 0, \lambda > 1, k_{\min} - \frac{1}{2\tau} - \frac{1}{2\lambda} k_{\max} > 0, \rho > 0 \) is designed to ensure that \( \| e \|^2 > \sqrt{\rho} > 0 \), i.e., the minimum positive triggering interval exists and the event generator is not triggered infinitely. Thus, the Zeno behavior can be avoided.

**V. STABILITY ANALYSIS OF CLOSED-LOOP SYSTEM**

The system states are transmitted to the controller in a package-based manner. Under the event-triggered modular, the transmissions can only occur at discrete instants \( t_i \), \( s \in Z^+ \), satisfying \( 0 \leq t_1 < t_2 < \ldots \). Hence, the closed-loop system is modeled as an impulsive system with a flow dynamic during the flow interval \( t \in (t_i, t_{i+1}) \) and a jump dynamic at time instant \( t = t_i \). The augmented state vector is defined as:

\[
\xi = \begin{bmatrix}
Z^T & \hat{Z}^T & \hat{W}_c^T
\end{bmatrix}^T, \quad \Phi(\hat{W}_c)
\]

where the expression of \( \Phi(\hat{W}_c) \) is given in Eqns.(46), \( C \) is the flow set defined by event-triggered conditions (47).

Similarly, the reset dynamics can be deduced as:

\[
\Delta \xi = \begin{bmatrix}
0 \\
e(t)
\end{bmatrix}, \quad \xi \notin C
\]
**Theorem 1:** For the nonlinear strict-feedback system (1), providing the Assumptions 1-3 are satisfied, the feedforward controller is designed as (7), the feedback control-disturbance pairs are renewed by the policy (39), the performance cost function are approximated by critic NN with the weight tuning law (45), and the system state and controller are updated by the triggered conditions (47). The resulting nonlinear impulsive closed-loop system is asymptotically stable updated by the triggered conditions (47).

Proof: Case 1 (During the flow period $t_s \leq t < t_{s+1}$, i.e., events are not triggered): Selecting the Lyapunov function candidate as:

$$V_c(t) = V_c(t) + V_c(t) + V_{\bar{W}}(t)$$

where

$$V_c(t) = \sum_{i=1}^{n} \frac{z_i(t)}{2},$$

$$V_{\bar{W}}(t) = \frac{1}{2} \bar{W}_c(t) \tilde{c}^{-1} \bar{W}_c(t) + J_s(Z).$$

The first-order time derivative of the first term $V_c(t)$ can be expressed as:

$$\dot{V}_c \leq \sum_{i=1}^{n} \left[ -\left( k_i - \frac{1}{2} g_{i-1} \right) \bar{z}_i \tilde{c}_i + k_i \zeta_i e_i \right] + Z^T H(Z) + Z^T G_a U^a + Z^T G_d D$$

Considering Assumption 2, yields that:

$$Z^T (H(Z) + G_a U^a + G_d D) \leq \frac{1}{2\tau} \|Z\|^2 + \frac{\tau c^2}{2} \|Z\|$$

The first-order time derivative of the second term $V_c(t)$ is:

$$\dot{V}_c = 0$$

The third term $V_{\bar{W}}(t)$ satisfies that:

$$\dot{V}_{\bar{W}}(t) = \bar{W}_c(t) \bar{c}^{-1} \bar{W}_c(t) + (\nabla J_s(Z))^T \dot{Z}$$

Recalling the event triggering condition (47), the Eqs.(59) satisfies:

$$\dot{V}_c(t) \leq -\sigma (k_{\min} - \frac{1}{2\tau} - \frac{1}{2\lambda} k_{\max}) \|Z\|^2$$

$$- (1 - \sigma) (k_{\min} - \frac{1}{2\tau} - \frac{1}{2\lambda} k_{\max}) \|Z\|^2$$

$$+ \frac{1}{2} \lambda k_{\max} \|e\|^2 + \frac{\tau c^2}{2} \|Z\| - \lambda_{\min}(K) \|\Upsilon\|^2 + T \|\Upsilon\|^2$$

Note that $\bar{W}_c = W_c - \bar{W}_c$, the following formula holds:

$$\bar{W}_c(t) \left( \kappa_1 W_c - \kappa_2 u_1^T \bar{W}_c \right)$$

$$= \bar{W}_c(t) \kappa_1 W_c - \bar{W}_c(t) \kappa_1 \bar{W}_c$$

$$- \bar{W}_c(t) \kappa_2 u_1^T W_c + \bar{W}_c(t) \kappa_2 u_1^T \bar{W}_c$$

(55)

Considering (34) and (39), we have

$$(U^a)^* - \hat{U}_a = - \frac{1}{2} G_a \left( \nabla \phi_c^T(\hat{Z}) \bar{W}_c + \nabla e_c(\hat{Z}) \right)$$

$$D^* - \hat{D} = \frac{1}{2} \eta^2 G_d \left( \nabla \phi_c^T(\hat{Z}) \bar{W}_c + \nabla e_c(\hat{Z}) \hat{Z} \right)$$

According to Lemma 1, we can deduce that, (56), as shown at the bottom of the next page, where (57), as shown at the bottom of the next page.

The parameters $\lambda_1, \lambda_2$ are tuned to ensure the matrix $K$ is positive definite, and Eqs.(56) is:

$$\dot{V}_c(t) \leq -(k_{\min} - \frac{1}{2\tau}) \|Z\|^2 + k_{\max} \|Z\| \|e\|$$

$$+ \frac{\tau c^2}{2} \|Z\| - \lambda_{\min}(K) \|\Upsilon\|^2 + \|T\| \|\Upsilon\|$$

$$\leq -(k_{\min} - \frac{1}{2\tau} - \frac{1}{2\lambda} k_{\max}) \|Z\|^2 + \frac{1}{2} \lambda k_{\max} \|e\|^2$$

$$+ \frac{\tau c^2}{2} \|Z\| - \lambda_{\min}(K) \|\Upsilon\|^2 + \|T\| \|\Upsilon\|$$

(58)

It can be concluded that $\|T\|$ is upper bounded $\|T\| \leq T$, supported by Assumption 3. Then the following inequation holds:

$$\dot{V}_c(t) \leq -\sigma (k_{\min} - \frac{1}{2\tau} - \frac{1}{2\lambda} k_{\max}) \|Z\|^2$$

$$- (1 - \sigma) (k_{\min} - \frac{1}{2\tau} - \frac{1}{2\lambda} k_{\max}) \|Z\|^2$$

$$+ \frac{1}{2} \lambda k_{\max} \|e\|^2 + \frac{\tau c^2}{2} \|Z\| - \lambda_{\min}(K) \|\Upsilon\|^2 + T \|\Upsilon\|$$

(59)

as long as the following conditions hold:

$$\|Z\| \geq \sqrt{\frac{\nu}{\sigma(k_{\min} - \frac{1}{2\tau} - \frac{1}{2\lambda} k_{\max})}} + \frac{\tau c^2}{4 \lambda_{\min}(K)} + \frac{T^2}{16 \sigma (k_{\min} - \frac{1}{2\tau} - \frac{1}{2\lambda} k_{\max}) + 4 \lambda_{\min}(K)} < 0$$

(60)
\[ \|\dot{Y}\| \geq \sqrt{\frac{v}{\lambda_{\min}(K)}} + \frac{T}{4\lambda_{\min}(K)} \left( \tau c_{\ell}^2 \right) + \frac{T^2}{4\lambda_{\min}(K)} \]  

\[ v = \frac{16\sigma(k_{\min} - \frac{1}{2\tau} - \frac{1}{2\tau} k_{\max})}{2\tau^2} + \frac{T^2}{4\lambda_{\min}(K)} \tag{61} \]

Therefore, all the signals in the closed-loop system realize UUB according to Lyapunov theorem during the flow \( t_s \leq t < t_{s+1} \).

**Case 2** (At the jump instants \( t = t_s \), i.e., events are triggered): Choosing the same Lyapunov function candidate (50), and the first difference of the first term and second term is:

\[ \Delta V_c = \sum_{i=1}^{n} \frac{1}{2} \left( \dot{\xi}_i^2 \right) - \sum_{i=1}^{n} \frac{1}{2} \left( \ddot{\xi}_i \right)^2 \equiv 0, \quad t = t_s \]

\[ \Delta V_c = \sum_{i=1}^{n} \frac{1}{2} \left( \dot{\xi}_i^2 \right) - \sum_{i=1}^{n} \frac{1}{2} \left( \ddot{\xi}_i \right)^2 \]

\[ \leq \frac{1}{2} (d^2)^2 - \sum_{i=1}^{n} \frac{1}{2} \left( \ddot{\xi}_i \right)^2, \quad t = t_s \]  

We can conclude that \( \dot{V}_c(t) < 0 \) for \( t \in [t_s, t_{s+1}) \), which indicates that \( V_c(t) \) is monotonically decreasing. Then the first
difference of the third term is:
\[
\Delta V_c = \left( \frac{1}{2} (\hat{W}_c^+)^T l_c^{-1} \hat{W}_c^+ + J_c(Z^+) \right) - \left( \frac{1}{2} (\hat{W}_c)^T l_c^{-1} \hat{W}_c + J_c(Z) \right) \leq 0 \quad (63)
\]

Finally, the first difference of the Lyapunov function (50) is:
\[
\Delta V_c \leq \frac{1}{2} (D_1^2) - \sum_{i=1}^{n} \frac{1}{2} (\hat{z}_i)^2 \quad (64)
\]

which satisfies that \( \Delta V_c < 0 \) as long as \( \| \hat{Z} \| > D_1^2 \). Therefore, the system tracking errors \( Z \), the event-referred tracking errors \( \hat{Z} \), and the critic NN weight errors \( \hat{W}_c(t) \) are bounded at the jump instants.

The analysis for the two cases renders that the close-loop nonlinear impulsive system state \( z \) is semi-globally ultimately bounded under the event triggered condition (47). \( \square \)

VI. NUMERICAL SIMULATIONS

**Example 1:** A second-order strict-feedback system is considered as a numerical example.
\[
\begin{align*}
\dot{x}_1 &= -0.1x_1^3 + x_2 + d_1 \\
\dot{x}_2 &= -0.5x_1 + 0.5x_2 - x_1^2x_2 + u + d_2 \\
y &= x_1
\end{align*}
\]
where \( f_1(x_1) = -0.1x_1^3, g_1 = 1, f_2(x) = -0.5x_1 + 0.5x_2 - x_1^2x_2, g_2 = 1, d_1, d_2 \) are the unknown disturbances.

The reference signal is chosen as \( x_{id}(t) = \sin(t) + 0.5 \sin(2t) \). The system state vector is initialized as \( x(0) = [1, 0]^T \). For the numerical example (65), the feedforward controller and the feedback control-disturbance pairs are designed as:
\[
\begin{align*}
\begin{cases}
\begin{aligned}
\bar{u}_1^b &= g_1^{-1} (-k_1 \hat{z}_1 - f_1(x_{id}) + \dot{x}_{id}) \\
\bar{u}_2^b &= g_2^{-1} (-k_2 \hat{z}_2 - f_2(\hat{x}_{2id})) 
\end{aligned}
\end{cases}
\end{align*}
\]
\[
\begin{align*}
\hat{U}(\hat{Z}) &= -\frac{1}{2} \Xi^{-1} G_u^T (\nabla \psi_c^T (\hat{Z}) \hat{W}_c(t)) \\
\hat{D}(Z) &= \frac{1}{2n^2} G_d^T (\nabla \psi_c^T (Z) \hat{W}_c(t))
\end{align*}
\]

The gains for feedforward tracking controller are selected as \( k_1 = 2, k_2 = 3 \). For the critic network in optimal feedback controller, the activation function is experimentally selected as \( \phi_c(Z) = [z_1, z_2, z_1 \tan^{-1}(z_1), z_2 \tan^{-1}(z_2), z_1^2, z_2^2, z_1^3, z_2^3]^T \). The initial weight values are chosen randomly in the interval \([-1, 1]\). In order to ensure the convergence of critic weights, the persistence excitation signal \( \sin^3(t) \cos(t) + \sin^3(2t) \cos(2t) \) is added to the control input during the first 5s. The parameters in the utility function are chosen as \( Q = \text{diag} [10, 10]^T, \Xi = \text{diag} [20, 10]^T, \eta = 4 \).

The remaining parameters are configured as:
\[
\begin{align*}
\nabla J_c(Z) &= 0.5(z_1^2 + z_2^2), & l_c &= 20, & \lambda_1 &= 200, & \lambda_2 &= 20, & \sigma &= 0.01, & \tau &= 1, & \lambda &= 4, & \rho &= 0.05
\end{align*}
\]

System output trajectory in Fig. 1 exhibits the output can track the reference signal within an acceptable tracking error. The evolution of other system state and control input during the learning phase are shown in Fig.2 and Fig.3. The event sampling instants in Fig.4 are changing and non-periodic. The trajectories of estimated critic weights in Fig.5 indicate that the critic networks weights can converge to the optimal values gradually and the optimal controller is successfully approximated.

**FIGURE 1.** Trajectory of the system output.

**FIGURE 2.** Evolution of other system state.

**FIGURE 3.** Evolution of control input.

**Example 2:** A single-link robot manipulator with a motor rotor is considered as a practical example.
\[
\begin{align*}
\dot{x}_1 &= x_2 \\
\dot{x}_2 &= -N \sin(x_1)/M - Bx_2/M + x_3/M + d_2 \\
\dot{x}_3 &= -K_B/L - Rx_3/L + u/L + d_3
\end{align*}
\]
\[ (67) \]
where $x_1$ is the motor angular position, $x_2$ is the motor angular rate, $x_3$ is the motor armature current, $u$ is the control voltage input, $d_2$, $d_3$ is disturbance,

$$
N = (mL_0g/2 + M_0L_0g)/K_T
M = (J + mL_0^2/3 + M_0L_0^2 + 2M_0R_0^2/5)/K_T
B = B_0/K_T
$$

The physical parameters are: the link mass $m = 0.506$, the link length $L_0 = 0.305$, gravity coefficient $g = 9.81$, the load mass $M_0 = 0.434$, electromechanical conversion coefficient $K_T = 0.9$, the load radius $R_0 = 0.023$, viscous friction coefficient $B_0 = 0.01625$, the rotor inertia $J = 0.001625$, the back EMF coefficient $K_R = 0.9$, the armature inductance and resistance $L = 15$, $R = 5$. The reference signal is $x_{1d}(t) = \sin(t)$. The initial system state is $x(0) = [0.1, 0.5, 0.5]^T$.

For the practical example (67), the feedforward controller and the feedback controller are proposed as:

$$
\begin{cases}
   u_1^b = g_1^{-1}(-k_1z_1 - f_1(x_{1d}) + \dot{x}_{1d}) \\
   u_2^b = g_2^{-1}(-k_2z_2 - f_2(\dot{x}_{2d})) \\
   u_3^b = g_3^{-1}(-k_3z_3 - f_3(\dot{x}_{3d})) \\
   \hat{\dot{\theta}}(Z) = \frac{1}{2}Z^{-1}G_u^T (\nabla \varphi_C^T(\hat{Z})\hat{W}_e(t)) \\
   \dot{D}(Z) = \frac{1}{2\eta^2}G_d^T (\nabla \varphi_C^T(Z)\hat{W}_e(t))
\end{cases}
$$

The feedforward controller gains are $k_1 = 1$, $k_2 = 3$, $k_3 = 3$. In optimal feedback controller, the activation function in critic network is chosen as $\varphi(Z) = [z_1, z_2, z_3, z_1 \tan^{-1}(z_1), z_2 \tan^{-1}(z_2), z_3 \tan^{-1}(z_3), z_1^3, z_2^3, z_3^3]^T \in \mathbb{R}^9$, which implies the critic network has nine neurons. The initial weight values are selected randomly in the interval $[-1, 1]$. The persistence excitation signal is utilized as same as Example 1. The parameters in the utility function are $Q = diag[10, 10, 10]^T$, $\Xi = diag[20, 10, 10]^T$, $\eta = 4$. The remaining parameters are:

$$
\nabla J_1(Z) = 0.5(z_1^2 + z_2^2 + z_3^2), \quad l_c = 20, \quad \lambda_1 = 200,
\lambda_2 = 20, \quad \sigma = 0.01, \quad \tau = 1, \quad \lambda = 4, \quad \rho = 0.05
$$

Evolution of the angular position, angular rate and armature current are depicted in Fig.6-Fig.8. The simulation fig shows that the angular position can track to the reference signal despite of the transient behavior during the first 5s when the persistence excitation signal is turned on. The event sampling instants in Fig.9 show that less sampling instants are required than the time-triggered control. The proposed event driven control can save communication resources and control costs greatly.
Note that the trajectories of estimated critic weights plotted in Fig.5 and Fig.10 converge near zero, as the backstepping-based feedforward control works as a main controller, and the ADP-based optimal feedback control works as an auxiliary controller. The optimal controller plays an important role during the transient process in Fig.5 and Fig.10 (it is more obvious in Fig.10). From a theoretical perspective, the optimal controller is constructed to stabilize the tracking errors in backstepping control, which converge to zero. Hence, the critic weights converge near zero.

From the simulation results for Example 1 and Example 2, the proposed optimal control can guarantee the stability of the closed-loop system and provide an excellent control performance in an optimal manner, while limited samples and fewer transmissions are performed in the event driven environment.

VII. CONCLUSION
This paper introduced an adaptive optimal control for nonlinear strict-feedback systems with generalized ET-learning scheme. To reduce calculations and network communication, an event generator is devised. A suitable feedforward control is proposed by the backstepping method; while the optimal feedback control is designed by the ADP technique to stabilize the tracking error dynamics. For the optimal feedback control, a critic network is constructed to approximate the saddle-point equilibrium in the two-player zero sum differential games. Due to the event triggering mechanism, the resulting closed-loop system fall into the impulsive closed-loop system form. The proposed optimal control strategy not only ensures that all signals in the impulsive closed-loop system are bounded, but also guarantees the predefined cost function is minimized. Finally, two numerical examples demonstrate the performance of the proposed control algorithm.

In the future, the adaptive optimal control problems for large-scale systems or multi-agent systems in strict-feedback form is worthy of further investigation.

REFERENCES

[1] B. Xu, D. Yang, Z. Shi, Y. Pan, B. Chen, and F. Sun, “Online recorded data-based composite neural control of strict-feedback systems with application to hypersonic flight dynamics,” IEEE Trans. Neural Netw. Learn. Syst., vol. 29, no. 8, pp. 3839–3849, Aug. 2018.
[2] T. Wang, J. Wu, Y. Wang, and M. Ma, “Adaptive fuzzy tracking control for a class of strict-feedback nonlinear systems with time-varying input delay and full state constraints,” IEEE Trans. Fuzzy Syst., early access, Nov. 11, 2019, doi: 10.1109/TFUZZ.2019.2952832.
[3] Y. Ji, W. Chi, and B. Tian, “Singular perturbation based command filtered back-stepping control for strict-feedback system and application to quadrotor,” in Proc. Chin. Automat. Cong. (CAC), Jinan, China, 2017, pp. 1394–1400.
[4] Y. Zou and Z. Zheng, “A robust adaptive RBFNN augmenting backstepping control approach for a model-scaled helicopter,” IEEE Trans. Control Syst. Technol., vol. 23, no. 6, pp. 2344–2352, Nov. 2015.
[5] B. Cui, Y. Xia, K. Liu, and G. Shen, “Finite-time tracking control for a class of uncertain strict-feedback nonlinear systems with state constraints: A smooth control approach,” IEEE Trans. Neural Netw. Learn. Syst., early access, Jan. 8, 2020, doi: 10.1109/TNNLS.2019.2959016.
[6] M. Van, M. Mavrocvitovitis, and S. S. Ge, “An adaptive backstepping nonsingular fast terminal sliding mode control for robust fault tolerant control of robot manipulators,” IEEE Trans. Syst., Man, Cybern., Syst., vol. 49, no. 7, pp. 1448–1458, Jul. 2019.
[7] C. L. P. Chen, G. X. Wen, Y. J. Liu, and Z. Liu, “Observer-based adaptive backstepping consensus tracking control for high-order nonlinear semi-strict-feedback multiagent systems,” IEEE Trans. Cybern., vol. 46, no. 7, pp. 1591–1601, Jul. 2016.
[8] Z. Ren, B. Zhao, and D. T. Nguyen, “Finite-time backstepping of a nonlinear system in strict-feedback form: Proved by Bernoulli inequality,” IEEE Access, vol. 8, pp. 47768–47775, 2020.
[9] J. Ma, S. Xu, G. Cui, W. Chen, and Z. Zhang, “Adaptive backstepping control for strict-feedback non-linear systems with input delay and disturbances,” IET Control Theory Appl., 13, no. 4, pp. 506–516, 2019.
[10] A. S. Malik, I. Ahmad, A. U. Rahman, and Y. Islam, “Integral backstepping and nonsingular control of magnetic levitation system,” IEEE Access, vol. 7, pp. 173230–173239, 2019.
[11] Q. Chen, H. Shi, and M. Sun, “Echo state network-based backstepping adaptive iterative learning control for strict-feedback systems: An error-tracking approach,” IEEE Trans. Cybern., early access, Aug. 16, 2019, doi: 10.1109/TCYB.2019.2931877.
[12] J. Huang, T. Zhang, Y. Fan, and J.-Q. Sun, “Control of rotary inverted pendulum using model-free backstepping technique,” IEEE Access, vol. 7, pp. 96965–96973, 2019.
[13] Z.-P. Jiang and Y. Jiang, “Robust adaptive dynamic programming for linear and nonlinear systems: An overview,” Eur. J. Control, vol. 19, no. 5, pp. 417–425, 2013.
[14] N. Malla and Z. Ni, “A new history experience replay design for model-free adaptive dynamic programming,” Neurocomputing, vol. 266, pp. 141–149, Nov. 2017, doi: 10.1016/j.neucom.2017.04.069.
[15] W. Bai, T. Li, and S. Tong, “NN reinforcement learning adaptive control for a class of nonstrict-feedback discrete-time systems,” IEEE Trans. Cybern., early access, Jan. 24, 2020, doi: 10.1109/TCYB.2020.2963849.
[16] Y. Liu, H. Zhang, R. Yu, and Z. Xing, “H∞ tracking control of discrete-time system with delays via data-based adaptive dynamic programming,” IEEE Trans. Syst., Man, Cybern., Syst., early access, Oct. 23, 2019, doi: 10.1109/TSMC.2019.2946397.
[17] M.-B. Radae and T. Lala, “Learning output reference model tracking for higher-order nonlinear systems with unknown dynamics,” Algorithms, vol. 12, no. 6, p. 212, 2019, doi: 10.3390/ai2100212.
[18] Z. Wang, R. Lu, F. Gao, and D. Liu, “An indirect data-driven method for trajectory tracking control of a class of nonlinear discrete-time systems,” IEEE Trans. Ind. Electron., vol. 64, no. 5, pp. 4121–4129, May 2016, doi: 10.1109/TIE.2016.2617830.
[19] H. Zargarzadeh, T. Dierks, and S. Jagannathan, “Optimal control of nonlinear continuous-time systems in strict-feedback form,” IEEE Trans. Neural Netw. Learn. Syst., vol. 26, no. 10, pp. 2535–2549, Jun. 2015.

[20] S. Tong, K. Sun, and S. Sui, “Observer-based adaptive fuzzy decentralized optimal control design for strict-feedback nonlinear large-scale systems,” IEEE Trans. Fuzzy Syst., vol. 26, no. 2, pp. 569–584, Apr. 2018.

[21] J. Sun and C. Liu, “Distributed fuzzy adaptive backstepping optimal control for nonlinear multimissile guidance systems with input saturation,” IEEE Trans. Fuzzy Syst., vol. 27, no. 3, pp. 447–461, Mar. 2019.

[22] H. Zhang, J. Zhang, G.-H. Yang, and Y. Luo, “Leader-based optimal coordination control for the consensus problem of multiagent differential games via fuzzy adaptive dynamic programming,” IEEE Trans. Fuzzy Syst., vol. 23, no. 1, pp. 152–163, Feb. 2015.

[23] C. Mu, K. Wang, Z. Ni, and C. Sun, “Cooperative differential game-based optimal control and its application to power systems,” IEEE Trans. Ind. Informat., vol. 16, no. 8, pp. 5169–5179, Nov. 2020.

[24] J. Sun, C. Liu, and X. Zhao, “Backstepping-based zero-sum differential games for missile-target interception systems with input and output constraints,” IET Control Theory Appl., vol. 12, no. 2, pp. 243–253, Nov. 2017.

[25] D. Wang, C. Mu, D. Liu, and H. Ma, “On mixed data and event driven design for adaptive-critic-based nonlinear H∞ control,” IEEE Trans. Neural Netw. Learn. Syst., vol. 29, no. 4, pp. 993–1005, Feb. 2018.

[26] X. Su, Z. Liu, Y. Zhang, and C. L. P. Chen, “Event-triggered adaptive fuzzy tracking control for uncertain nonlinear systems preceded by unknown Prandtl-Ishlinskii hysteresis,” IEEE Trans. Cybern., early access, Nov. 12, 2019, doi: 10.1109/TCYB.2019.2949022.

[27] C. Wang, L. Guo, C. Wen, and Q. Hu, “Event-triggered adaptive attitude tracking control for spacecraft with unknown actuator faults,” IEEE Trans. Ind. Electron., vol. 67, no. 3, pp. 2241–2250, Mar. 2020.

[28] Y.-X. Li and G.-H. Yang, “Model-based adaptive event-triggered control of strict-feedback nonlinear systems,” IEEE Trans. Neural Netw. Learn. Syst., vol. 29, no. 4, pp. 1033–1045, Feb. 2018.

[29] Y.-X. Li and G.-H. Yang, “Event-triggered adaptive backstepping control for parametric strict-feedback nonlinear systems,” Int. J. Robust Nonlinear Control, vol. 28, pp. 976–1000, Feb. 2018.

[30] L. Cao, H. Li, and Q. Zhou, “Adaptive intelligent control for nonlinear strict-feedback systems with virtual control coefficients and uncertain disturbances based on event-triggered mechanism,” IEEE Trans. Cybern., vol. 48, no. 12, pp. 3390–3402, Dec. 2018.

[31] C. Mu and K. Wang, “Approximate-optimal control algorithm for constrained zero-sum differential games through event-triggering mechanism,” Nonlinear Dyn., vol. 95, pp. 2639–2657, Dec. 2018, doi: 10.1007/s11071-018-4713-0.

[32] S. Xue, B. Luo, and D. Liu, “Event-triggered adaptive dynamic programming for zero-sum game of partially unknown continuous-time nonlinear systems,” IEEE Trans. Syst., Man, Cybern., Syst., early access, Jul. 17, 2018, doi: 10.1109/TSMC.2018.2852810.