Event-Based Robust Control for Uncertain Nonlinear Systems Using Adaptive Dynamic Programming

Qichao Zhang, Student Member, IEEE, Dongbin Zhao, Senior Member, IEEE, and Ding Wang, Member, IEEE

Abstract—In this paper, the robust control problem for a class of continuous-time nonlinear system with unmatched uncertainties is investigated using an event-based control method. First, the robust control problem is transformed into a corresponding optimal control problem with an augmented control and an appropriate cost function. Under the event-based mechanism, we prove that the solution of the optimal control problem can asymptotically stabilize the uncertain system with an adaptive triggering condition. That is, the designed event-based controller is robust to the original uncertain system. Note that the event-based controller is updated only when the triggering condition is satisfied, which can save the communication resources between the plant and the controller. Then, a single network adaptive dynamic programming structure with experience replay technique is constructed to approach the optimal control policies. The stability of the closed-loop system with the event-based control policy and the augmented control policy is analyzed using the Lyapunov approach. Furthermore, we prove that the minimal intersample time is bounded by a nonzero positive constant, which excludes Zeno behavior during the learning process. Finally, two simulation examples are provided to demonstrate the effectiveness of the proposed control scheme.

Index Terms—Adaptive dynamic programming (ADP), event-based control, neural network (NN), robust control, unmatched uncertainties.

I. INTRODUCTION

A S MANY practical control systems become more and more complex, uncertainties arise in the system models frequently. There are uncertainties associated with the modeling of dynamical systems, namely, plant uncertainties that can be structured (parametric uncertainties) and unstructured. There are uncertainties associated with the operation of dynamical systems related to exogenous disturbances [1] usually captured via disturbance inputs. These uncertainties may severely degrade the system performance, and even lead to system instability, so it is necessary to design robust controllers for uncertain nonlinear systems. According to whether the exogenous disturbances are in the range space of the input matrix, the uncertainties are often divided into matched and unmatched ones. Such a robust control problem of the systems with matched or unmatched uncertainties has been investigated recently by many approaches, such as $H_{\infty}$ approach [2], model predictive approach [3], sliding model approach [4], and so on. Lin et al. [5] established a connection between the robust control problem and the optimal control problem for the nonlinear systems with matched uncertainties by designing a corresponding optimal controller. Later, the connection between the optimal control problem and robust stabilization has been studied widely in recent years. In [6], an optimal controller was designed by solving the algebraic Riccati equation to guarantee linear robust stabilization of robot manipulators with matched and unmatched uncertainties. Lin [7] proved that the optimal controller can stabilize the linear and nonlinear uncertain systems in the same way. However, the detailed approach to solve the Hamilton–Jacobi–Bellman (HJB) equation for the nonlinear systems was not discussed.

As is known, it is intractable to give an analytic solution to the HJB equation for the nonlinear systems [8]. Recently, adaptive dynamic programming (ADP), which was proposed by Werbos [9], has been widely applied to approximate the solution of the HJB equation. For example, the $H_{\infty}$ control approach based on ADP was investigated for the uncertain nonlinear systems in [10]. Jiang and Jiang [11] proposed a robust ADP methodology for nonlinear uncertain systems. This methodology was applied to the power systems in [12]. Liu et al. [13] investigated the optimal robust guaranteed cost control problem and transformed this problem into a corresponding optimal control problem with an appropriate cost function, where an initial stabilizing control was not required for the proposed ADP algorithm. In [14], the robust control problem of nonlinear systems with matched uncertainties was converted into an optimal control problem of nominal systems. The robust controller was derived by adding a feedback gain to
the obtained optimal controller. In [15], the robust controller of nonlinear systems with unmatched uncertainties was designed by finding the optimal controller based on ADP with actor-critic structure. For the nonlinear deterministic systems, Zhao and Zhu [16] and Zhao et al. [17] developed several ADP methods to solve the corresponding optimal control problems. In [18], a supervised ADP was applied to the adaptive cruise control (ACC) of the automobile. However, the aforementioned approaches are conducted predicated on the traditional time-triggered strategy.

In general, the amount of transmitted data are huge using the traditional time-triggered approach, which is impractical for many resource-limited systems. Furthermore, the huge number of transmitted data often leads to tremendous computational burden. To mitigate the unnecessary waste of communication resources [19], event-triggered control (ETC) method has received great interests among the control researchers. Based on the triggering mechanism, the ETC is divided into event-based control and self-triggered control in [20]. The former one is reactive which requires a continuous monitoring of the triggering condition to determine triggering instants, while the latter is proactive which can precompute the next triggering instant using the knowledge of the plant dynamics to avoid the continuous monitoring. ETC has been widely used in the distributed networked control systems [21], complex dynamical networks [22]–[24], and multiagent systems [25]–[28] to solve the consensus problem. Recently, the event-based control has been integrated with the ADP approach to solve the optimal control problems. Sahoo et al. [29] proposed a neural network (NN)-based ETC scheme for nonlinear discrete-time systems using ADP approach. This scheme was extended to the unknown nonlinear continuous-time (CT) systems in [30]. Vamvoudakis [31] proposed an optimal adaptive ETC algorithm based on the actor-critic structure for CT nonlinear systems with guaranteed performance. In [32], the event-based reinforcement learning approach was developed for the nonlinear systems without requiring the exact knowledge of system dynamics. However, to the best of our knowledge, system uncertainties are not addressed in the existing work on event-based optimal control. Evidently, it is more difficult to design the triggering condition due to the existence of uncertainties. That is, the lack of nonlinear robust stabilization via event-based control and ADP technique motivates our research.

In this paper, we investigate the robust control problem of nonlinear systems with unmatched uncertainties using an optimal control approach. First of all, the controlled problem is transformed into an optimal control problem of an auxiliary system with a newly modified cost function. Then, an adaptive triggering condition is derived to guarantee the stability of the uncertain system under the event-based optimal controller. In order to obtain the optimal control policy, the ADP technique with a single critic network structure is used to solve the corresponding HJB equation. Finally, an event-based ADP algorithm is presented by designing the NN weight updating law based on experience replay technique. Another adaptive triggering condition during the NN learning process is derived which can guarantee the uniform ultimate boundedness (UUB) of critic NN’s error dynamics and the stability of the closed-loop impulsive system. Compared with some existing literature, the main contributions emphasize in two parts.

1) Motivated by [5], [6], and [13], we further consider the nonlinear robust control problem based on the ETC mechanism. Compared with the existing work of event-based optimal control in [29]–[32], the unmatched uncertainty is considered in this paper, which increases the difficulty to design the event-based control scheme. To compensate for the unmatched uncertainty, an augmented control input and a modified cost function are introduced. Accordingly, a novel adaptive triggering condition associated with the augmented control policy is derived to guarantee the robust stabilization.

2) Combined with the experience replay technique in [17], the proposed event-based ADP algorithm can approach the optimal control policies with a relaxed persistence of excitation (PE) condition. Different from the actor-critic structure in [29] and [31], a single critic NN structure is built in the proposed algorithm, which can reduce the computational burden. In addition, a positive constant bound of the minimal intersample time is given, which excludes the Zeno behavior.

The rest of this paper is organized as follows. Section II introduces the robust control problem of the nonlinear system with unmatched uncertainties and the traditional optimal control problem. In Section III, the connection between the robust stabilization and the optimal control problem is discussed under the event-based mechanism. In Section IV, the event-based ADP algorithm is proposed to approximate the optimal control policy with a stability analysis. Simulation results are shown in Section V. The conclusion and discussion is presented in Section VI.

II. PROBLEM STATEMENT

Consider the CT uncertain nonlinear system given by

\[ \dot{x}(t) = f(x(t)) + g(x(t))u(x) + k(x(t))W(x(t)) \]  

(1)

where \( x = x(t) \subseteq \mathbb{R}^n \) is the state vector, \( u = u(x) \subseteq \mathbb{R}^b \) is the control input, \( f(\cdot) \in \mathbb{R}^n, g(\cdot) \in \mathbb{R}^{n \times b}, \) and \( k(\cdot) \in \mathbb{R}^{n \times q} \) are the differentiable nonlinear dynamics with \( f(0) = 0 \), and \( W(\cdot) \in \mathbb{R}^q \) is the unknown nonlinear perturbation due to parameter variations and external disturbances. Assume that \( W(0) = 0 \), so that \( x = 0 \) is an equilibrium of system (1). The uncertainty \( W(x) \) is known as an unmatched uncertainty for system (1), if \( k(x) \neq g(x) \). Here, we provide Assumptions 1 and 2 for system (1) and the state-dependent perturbation \( W(x) \).

**Assumption 1:** \( f(x) \) and \( g(x) \) are both Lipschitz continuous on a compact set \( \Omega \subseteq \mathbb{R}^n \) such that \( \|f(x)\| \leq \ell_f \|x\| \) and \( \|g(x)\| \leq \ell_g \|x\| \).

**Assumption 2:**

1) The uncertainty \( W(x) \) is bounded by a known nonnegative function \( W_M(x) \), i.e., \( \|W(x)\| \leq W_M(x) \) with \( W_M(0) = 0 \).
2) There exists a nonnegative function \( g_M(x) \) such that
\[
\| g^+(x)k(x)W(x) \|^2 \leq \frac{g_M^2(x)}{2}
\]
where \( g^+(x) \) denotes the pseudoinverse of function \( g(x) \).

In this paper, we aim to find a control policy, so that system (1) is globally asymptotically stable for all unmatched uncertainties \( W(x) \). Note that Assumption 2 is general for the robust control problem and it can also be found in [15].

Motivated by [7], the robust control problem of the uncertain nonlinear system will be converted into an optimal control problem in the following.

First, the uncertain term \( k(x)W(x) \) is decomposed into a matched component and an unmatched one in the range space of \( g(x) \), that is
\[
k(x)W(x) = g(x)g^+(x)k(x)W(x) + (I - g(x)g^+(x))k(x)W(x).
\]

Then, we can transform the robust control problem into an optimal control problem as follows.

A. Optimal Control Problem

For the corresponding auxiliary system
\[
\dot{x} = f(x) + g(x)u(x) + (I - g(x)g^+(x))k(x)w(x) \quad (2)
\]
where \( w = w(x) \in \mathbb{R}^d \) is an augmented control to deal with the unmatched uncertainty component, and \([u^T(x), w^T(x)]^T\) is a control policy pair of system (2).

Assume that the auxiliary system (2) is controllable. It is desired to find the optimal control policy pair \([u^*(T), w^*(T)]^T\) that minimizes the cost function given by
\[
V(x(0)) = \int_0^\infty \| r^T \|^2 g_M^2(x) + \eta^2 \| m^T \|^2 W_M^2(x) + U(x, u, w) dt
\]
where the utility \( U(x, u, w) = x^T Q x + u^T(x)R u(x) + \eta^2 w^T(x)M w(x) \), and \( \eta > 0 \) is a designed parameter. Here, \( Q, R, \) and \( M \) are the positive definite symmetric matrices. According to the principle of Cholesky decomposition, we have \( R = r r^T \) and \( M = m m^T \), where \( r \) and \( m \) are two appropriate lower triangular matrices.

Remark 1: The cost function (3) is designed to reflect the modification related to the problem transformation. In fact, the optimal control policy related to this modified cost function can also be solved as a solution of the robust control problem. Compared with the classical one for the nominal system in [31], a term \( \| r^T \|^2 g_M^2(x) + \eta^2 \| m^T \|^2 W_M^2(x) \) is added in (3) to reflect the unmatched uncertainty of system (1). In contrast to that one in [14], the added term in (3) is not only related to the perturbation \( W(x) \) but related to the system drift dynamics, since the uncertainties are unmatched in this paper.

Remark 2: For the optimal control problem of the auxiliary system (2), the designed feedback control inputs should be admissible (see [33] for definition). In this paper, we use \( \Phi(\Omega) \) to denote the set of admissible policies on a compact set \( \Omega \).

For any admissible policies \( u, \omega \in \Phi(\Omega) \), if the cost function (3) is continuously differentiable, the infinitesimal version of (3) is the so-called nonlinear Lyapunov equation
\[
\nabla V^T(f(x) + g(x)u(t) + (I - g(x)g^+(x))k(x)w(t)) + \| r^T \|^2 g_M^2(x) + \eta^2 \| m^T \|^2 W_M^2(x) + U(x, u, w) = 0
\]
where \( \nabla V = \frac{\partial V}{\partial x} \) is the partial derivative of the cost function \( V(x) \) with respect to the state \( x \), and \( V(0) = 0 \).

Define the Hamiltonian function of system (2) as
\[
H(x, \nabla V, u, \omega) = \| r^T \|^2 g_M^2(x) + \eta^2 \| m^T \|^2 W_M^2(x) + U(x, u, w) + (\nabla V)^T f(x) + g(x)u + (I - g(x)g^+(x))k(x)w).
\]
The optimal cost function of system (2)
\[
V^*(x(0)) = \min_{u, \omega \in \Phi(\Omega)} \int_0^\infty \{ \| r^T \|^2 g_M^2(x(\tau)) + \eta^2 \| m^T \|^2 \times W_M^2(x(\tau)) + U(x(\tau), u(\tau), w(\tau)) \} d\tau
\]
satisfies the associated HJB equation
\[
\min_{u, \omega \in \Phi(\Omega)} H(x, \nabla V^*, u, \omega) = 0 \quad (5)
\]
where \( V^*(x) \) is a solution of the HJB equation.

Define \( d(x) = (I - g(x)g^+(x))k(x) \). Assume that the minimum policy pair on the left-hand side of (5) exists and is unique. According to the stationary conditions, the optimal control policies are given by
\[
\frac{\partial H}{\partial u}(x, \nabla V^*, u, \omega) = 0 \Rightarrow u^*(x) = -\frac{1}{2} R^{-1} g^T(x) \nabla V^*(x) \quad (6)
\]
\[
\frac{\partial H}{\partial \omega}(x, \nabla V^*, u, \omega) = 0 \Rightarrow w^*(x) = -\frac{1}{2\eta^2} M^{-1} d^T(x) \nabla V^*(x). \quad (7)
\]
Based on (6) and (7), the HJB equation (5) is rewritten as
\[
H(x, \nabla V^*, u^*, w^*) = (\nabla V^*)^T f(x) + x^T Q x + \| r^T \|^2 g_M^2(x) + \eta^2 \| m^T \|^2 W_M^2(x) - \frac{1}{4} (\nabla V^*)^T g(x) R^{-1} g^T(x) \nabla V^* - \frac{1}{4\eta^2}(\nabla V^*)^T d(x) M^{-1} d^T(x) \nabla V^* = 0. \quad (8)
\]

So far, the robust control problem is transformed into a corresponding time-triggered optimal control problem. Then, the traditional ADP technique can be employed to approximate the solution \( V^*(x) \) of the HJB equation. In order to reduce the computational burden and save communication resources, the ETC mechanism is introduced in this paper. An adaptive triggering condition is designed to guarantee the stability of the uncertain system with an event-based optimal controller.
III. EVENT-BASED ROBUST OPTIMAL CONTROLLER

To propose the ETC mechanism, we first define a monotonically increasing sequence of triggering instants \( \{ \tau_j \}_{j=0}^{\infty} \), where \( \tau_j \) is the \( j \)th consecutive sampling instant with \( \tau_j < \tau_{j+1} \), \( j \in \mathbb{N} \) with \( \mathbb{N} = \{0, 1, 2, \cdots \} \). Then, a sampled-data system characterized by the triggering instants is introduced, where the controller is updated based on the sampled state \( \hat{x}_j = x(\tau_j) \) for all \( t \in [\tau_j, \tau_{j+1}) \). Define the event-based error as

\[
e_j(t) = \hat{x}_j - x(t), \quad \forall t \in [\tau_j, \tau_{j+1}), \quad j \in \mathbb{N}
\]

where \( x(t) \) and \( \hat{x}_j \) denote the current state and the sampled state, respectively.

In the event-based control method, the triggering instants are determined by a triggering condition. In general, the triggering condition is determined by the event-based error and a designed state-dependent threshold. When the event-based error exceeds the state-dependent threshold, an event is triggered. Then, the system states are sampled, which resets the event-based error \( e_j(t) \) to zero. Accordingly, the designed event-based controller \( u(\hat{x}_j) = \mu(\hat{x}_j) \) is updated. Note that the system states are held until the next triggering instant. Clearly, the control signal \( \mu(\hat{x}_j) \) is a function of the event-based state vector, which is executed based on the latest sampled state \( \hat{x}_j \) instead of the current value \( x(t) \). That is, the event-based controller is only updated at the triggering instant sequence \( \{ \tau_j \}_{j=0}^{\infty} \) and remains unchanged in each time interval \( t \in [\tau_j, \tau_{j+1}) \). Hence, this control signal \( \mu(\hat{x}_j) \) with \( j \in \mathbb{N} \) is a piecewise constant function on each segment \( [\tau_j, \tau_{j+1}) \).

Under the event-triggered mechanism, the transformed optimal control problem in Section II can be restated as follows.

With the event-based control input \( \mu(\hat{x}_j) \), the sampled-data version of the auxiliary system (2) can be written as

\[
\dot{x}(t) = f(x) + g(x)\mu(x(t) + e_j(t)) + d(x)w(x(t)).
\]

Considering the event-based sampling rule, the optimal control policy (6) becomes

\[
\mu^*(\hat{x}_j) = -\frac{1}{2} R^{-1} g^T(\hat{x}_j) \nabla V^*(\hat{x}_j)
\]

for all \( t \in [\tau_j, \tau_{j+1}) \), where \( \nabla V^*(\hat{x}_j) = \partial V^*(x(\hat{x}_j) / \partial x) \mid_{x = \hat{x}_j} \).

By using the optimal cost function \( V^*(x) \), the event-based controller (11), and the augmented controller (7), the Hamiltonian function (4) becomes

\[
H(x, \nabla V^*, \mu^*(\hat{x}_j), w^*(x)) = (\nabla V^*)^T \dot{f}(x) + x^T Q x + \| R \|^2 \| \phi \|^2 S_M(x) + \eta^2 \| M^T \|^2 W_M(x)
\]

\[
- \frac{1}{2} g^T(\nabla V^*) g(x) R^{-1} g^T(\hat{x}_j) \nabla V^*(\hat{x}_j)
\]

\[
+ \frac{1}{4} g^T(\nabla V^*) g(x) R^{-1} g^T(\hat{x}_j) \nabla V^*(\hat{x}_j)
\]

\[
- \frac{1}{4} W_M(x) - d(x)W_M(x) + \nabla V^*(x).
\]

For convenience of analysis, Assumption 3 is introduced.

**Assumption 3:** The optimal controller \( u^*(x) \) is Lipschitz continuous with respect to the event-based error

\[
\| u^*(x) - u^*(\hat{x}_j) \| = \| u^*(x) - u^*(x + e_j) \| \leq L \| e_j \|
\]

where \( L \) is a positive real constant.

**Remark 3:** This assumption is satisfied in many applications where the controller are affine with respect to \( e_j \). Note that \( w(t) \) is not the direct control policy of the robust control system (1), but it plays an important role in finding the event-based optimal control policy \( \mu^*(\hat{x}_j) \) for system (10).

For the sampled-data system (10) with feedback control policy pair \( [\mu^T(\hat{x}_j), w^T(t)]^T \), we say that the system is input-to-state stability (ISS) with respect to the event-based error \( e_j(t) \) if there exists an ISS Lyapunov function for (10). The ISS Lyapunov function adopted from [34] is defined as follows.

**Definition 1 (ISS):** A smooth function \( V : \mathbb{R}^n \rightarrow \mathbb{R}^+ \) is said to be ISS stability Lyapunov function for the closed-loop system (10) if there is a class of \( \mathcal{K}_\infty \) function \( \psi, \psi, \psi, \) and \( \chi \) satisfying

\[
\psi(\|x\|) \leq V(x) \leq \psi^*(\|x\|)
\]

\[
\nabla V^T f(x) + g(x)x + d(x)w(x) \leq \chi(\|e_j(t)\|) - \psi^*(\|x\|)
\]

where \( \mathbb{R}^+ \) denotes the set \( x \in \mathbb{R}^n : x > 0 \).

We now study the connection of the controller (6) and the event-based controller (11) and derive Lemma 1.

**Lemma 1:** Suppose that Assumption 3 holds and \( V^* \) is an ISS Lyapunov function for the closed-loop system (10). The event-based optimal controller (11) is a discretized version of the traditional optimal controller (6).

**Proof:** Combine (8) and (12), we can obtain

\[
H(x, \nabla V^*, \mu^*(\hat{x}_j), w^*(x)) = (u^T x R u^*(x) - 2 u^T x R \mu^*(\hat{x}_j) + \mu^T(\hat{x}_j) R \mu^*(\hat{x}_j))
\]

\[
= (u^T x (u^*(x) - \mu^*(\hat{x}_j)))^T (u^T (u^*(x) - \mu^*(\hat{x}_j)))
\]

where \( u(\hat{x}_j) = \mu(\hat{x}_j) \).

Based on Assumption 3, we have

\[
H(x, \nabla V^*, \mu^*(\hat{x}_j), w^*(x)) = H(x, \nabla V^*, u^*(x), w^*(x)) \]

\[
\leq L^2 \| e_j(t) \|^2.
\]

Since \( V^* \) is an ISS Lyapunov function for (10), the closed-loop system is ISS with respect to \( e_j(t) \). Then, we have the state-dependent threshold \( \psi(\|x\|) \geq \chi(\|e_j(t)\|) \), and the event-based error \( e_j(t) \) is forced to be zero at the triggering instants. As the state asymptotically approaches to the origin, the state-dependent threshold becomes smaller. Consequently, the event-based optimal controller (11) can be regarded as a discretized version of (6).

**Remark 4:** Combine (8), (12), and (13), we have

\[
H(x, \nabla V^*, \mu^*(\hat{x}_j), w^*(x)) = (u^T x (u^*(x) - \mu^*(\hat{x}_j)) + d(x)w^*(x))
\]

\[
+ \| R \|^2 \| \phi \|^2 S_M(x) + \eta^2 \| M^T \|^2 W_M(x) + U(x, \mu^*(\hat{x}_j), w^*(x))
\]

\[
= (r^T (u^*(x) - \mu^*(\hat{x}_j)))^T (r^T (u^*(x) - \mu^*(\hat{x}_j))).
\]
function $H(x, \nabla V^*, \mu^*(\hat{x}_j), w^*(x))$ not equal to zero during each time interval $t \in (r_j, t_{j+1})$. Note that this error is not considered in [31] and [32]. Since the event-based controller is executed using the event-based state vector, a small transformation error means a higher number of events, which results in more transmissions and computations. Hence, a suitable triggering condition should be designed to attain a tradeoff between the robust stabilization performance and resource utilization.

Lemma 1 reveals the relationship between the optimal control problem (6) and the event-based controller (11). In order to solve the robust control problem using the event-based optimal controller (11), an adaptive triggering condition is presented in Theorem 1.

**Theorem 1:** Suppose that $V^*(x)$ is the solution of the HJB equation (8). For all $t \in [r_j, t_{j+1}], j \in \mathbb{N}$, the control policies are given by (7) and (11), respectively. If the triggering condition is defined as follows:

$$\|e_T(t)\|^2 > (1 - \beta^2)\frac{\|\mu^*(\hat{x}_j)\|^2}{2\|r^T\|^2\|L\|^2} \lambda_{\min}(Q)\|x\|^2 - \frac{\eta^2\|m^T w^*(x)\|^2}{\|r^T\|^2\|L\|^2}$$

the solution $\mu^*(\hat{x}_j)$ to the optimal control problem is also a solution to the robust control problem, where $e_T$ denotes the threshold, $\lambda_{\min}(Q)$ is the minimal eigenvalue of $Q$, and $\beta \in (0, 1)$ is a designed sample frequency parameter. That is, system (1) can be globally asymptotically stable for all uncertainties $W(x)$ under $\mu^*(\hat{x}_j)$.

**Proof:** Let $[\mu^T(\hat{x}_j), w^T(x)]$ and $V^*(x)$ be the optimal control policy pair and optimal cost function, respectively. According to the definition of cost function in (3), $V^*(x)$ can be seen as a Lyapunov function. Taking the derivative of $V^*(x)$ along the trajectory of system (1), we can obtain

$$\dot{V}^*(x) = (\nabla V^*)^T(f(x) + g(x)\mu^*(\hat{x}_j) + k(x)W(x))$$

$$= (\nabla V^*)^T(f(x) + g(x)\mu^*(\hat{x}_j) + g(x)^+ k(x)W(x))$$

$$+ (\nabla V^*)^T(I - g(x)g^+(x))k(x)W(x).$$

From (6) to (8), we can get

$$g^T(x)\nabla V^* = -2Ru^*(x),$$

$$d^T(x)\nabla V^* = -\eta^2Mw^*(x),$$

$$(\nabla V^*)^T f(x) = -x^TQx - \|r^T\|^2g^2 M^2(x) - \eta^2\|m^T\|^2W^2 M(x)$$

$$+ \frac{1}{4}(\nabla V^*)^T g(x)R^{-1}g^T(x)\nabla V^*$$

$$+ \frac{1}{4\eta^2}(\nabla V^*)^T d(x)M^{-1}d^T(x)\nabla V^*. \quad \text{(16)}$$

Substitute (17)–(19) into (16), we have

$$\dot{V}^*(x) = -x^TQx - \|r^T\|^2g^2 M^2(x) - \eta^2\|m^T\|^2W^2 M(x)$$

$$+ u^T(x)Ru^*(x) + \eta^2w^T(x)Mw^*(x)$$

$$- 2u^T(x)R\mu^*(\hat{x}_j) - 2u^T(x)Rg^+(x)k(x)W(x)$$

$$- 2\eta^2w^T(x)Mw(x). \quad \text{(19)}$$

Note that $R = r^TR$ and $M = mm^T$. By using the Lipschitz condition from Assumption 3, we can obtain

$$\dot{V}^*(x) \leq -x^TQx + 2\|r^T\|\|Le_j(t)\|^2 + 2\eta^2\|m^T w^*(x)\|^2$$

$$- \eta^2\|m^T w^*(x)\|^2 - \eta^2\|m^T w^*(x)\|^2$$

$$\leq \|r^T\|\|Le_j(t)\|^2 + 2\|r^T\|\|Le_j(t)\|^2 + 2\eta^2\|m^T w^*(x)\|^2.$$

Then, (20) can be rewritten as

$$\dot{V}^*(x) \leq -x^TQx + 2\|r^T\|^2\|Le_j(t)\|^2 + 2\eta^2\|m^T w^*(x)\|^2$$

$$- \eta^2\|m^T w^*(x)\|^2 - \eta^2\|m^T w^*(x)\|^2$$

$$\leq -x^TQx + 2\|r^T\|^2\|Le_j(t)\|^2 + 2\eta^2\|m^T w^*(x)\|^2.$$

According to Assumption 2, we have

$$\dot{V}^*(x) \leq -x^TQx + 2\|r^T\|^2\|Le_j(t)\|^2 + 2\eta^2\|m^T w^*(x)\|^2$$

$$= -x^TQx + 2\|r^T\|^2\|Le_j(t)\|^2 + 2\eta^2\|m^T w^*(x)\|^2 + 2\eta^2\|m^T w^*(x)\|^2 - 2\eta^2\|m^T w^*(x)\|^2.$$

If the triggering condition is defined as (15), we have

$$\dot{V}^*(x) \leq -\beta^2\lambda_{\min}(Q)\|x\|^2 < 0 \text{ for any } x(t) \neq 0 \text{ under the event-based control mechanism, that is the closed-loop system is asymptotically stable. The proof is completed.}$$

**Remark 5:** The sample frequency parameter $\beta$ and the designed parameter $\eta$ should be chosen to guarantee the term $\|e_T\|^2$ defined in (15) to be positive. For a fixed $\eta$, the sample frequency of sampled-data system (10) can be adjusted by the sample frequency parameter $\beta$ in the triggering condition (15). When $\beta$ is closer to 1, the system states are sampled more frequently, which means that the controller $\mu(\hat{x}_j)$ is updated more frequently.

**Remark 6:** Note that the control input $\mu^*(\hat{x}_j)$ is based on event-triggered mechanism, while the augmented control input $w^*(x)$ is based on time-triggered mechanism. The reasons why the augmented control is not based on event-triggered mechanism are as follows. On the one hand, the controller applied to the uncertain nonlinear system is the control input $\mu^*(\hat{x}_j)$ rather than the augmented control $w^*(x)$. On the other hand, it is difficult to derive the triggering condition (15), which can guarantee the robust stabilization, if the augmented control policy is also based on event-triggered method.
Remark 7: From Theorem 1, we know that the augmented control input $w^*(t)$ also plays an important role in acquiring the triggering condition (15) and guaranteeing the stability of the closed-loop system (1). When an event is triggered based on the triggering condition (15), the system states are sampled and the control policy (11) is updated accordingly. Then, the controller is held and utilized until the next triggering instant. Hence, the communication between the controller and the plant is reduced. It is significant for the situation that the controller has limited communication bandwidth.

IV. APPROXIMATE OPTIMAL CONTROLLER DESIGN

In this section, an online event-based ADP algorithm with a single NN structure is proposed to approximate the solution of the event-based HJB equation.

A. Event-Based ADP Algorithm via Critic Network

In the event-based ADP algorithm, only a single critic network with a three-layer network structure is required to approximate the optimal value function. The optimal value function based on NN can be formulated as

$$V^*(x) = W_c^T \phi(x) + \varepsilon$$

where $W_c \in \mathbb{R}^N$ is the critical NN ideal weights which in turn corresponds to the stabilizing solution of the underlying HJB, and not the ideal weights corresponding to other solutions of the underlying HJB. $\phi(x) \in \mathbb{R}^N$ is the activation function vector, $N$ is the number of hidden neurons, and $\varepsilon \in \mathbb{R}$ is the NN approximation error. We can obtain the partial derivative of (21) with respect to $x$ as

$$\nabla V^*(x) = \nabla \phi^T(x) W_c + \nabla \varepsilon.$$

Since the ideal weight matrix is unknown, the actual output of critic NN can be presented as

$$\hat{V}(x) = \hat{W}^T_c \phi(x)$$

where $\hat{W}_c$ represents the estimation of the unknown weight matrix $W_c$.

Accordingly, the augmented control policy (7) and the event-based control policy (11) can be approximated by

$$\hat{w}(x) = -\frac{1}{2\tau^2} M^{-1} d^T(x) \nabla \phi^T(x) \hat{W}_c$$

$$\hat{\mu}(\hat{x}_j) = -\frac{1}{2} R^{-1} g^T(\hat{x}_j) \nabla \phi^T(\hat{x}_j) \hat{W}_c.$$

Using the NN expression (21), the event-based HJB equation (14) becomes

$$H(x, W_c, \mu(\hat{x}_j), w)$$

$$= \| r^T \|^2 g_M^2(x) + \eta^2 \| m^T \|^2 W_M^2(x) + U(x, \mu(\hat{x}_j), w)$$

$$+ W_c^T \nabla \phi(f(x) + g(x)\mu(\hat{x}_j)) + d(x)w \right) \right)$$

$$= e_{cH} + (r^T (u(x) - \mu(\hat{x}_j)))^T (r^T (u(x) - \mu(\hat{x}_j)))$$

where $e_{cH}$ denotes the residual error. For fixed $N$, the NN approximation errors $e$ and $\nabla e$ are bounded locally [33]. That is, $\| \nabla e_{\text{max}} \|$, $\forall \nabla e_{\text{max}} : \sup \| \nabla e \| < \nabla e_{\text{max}}$. Assume that the residual error is bounded locally, which means there exists $e_{cH_{\text{max}}} > 0$ such that $|e_{cH}| \leq e_{cH_{\text{max}}}$.

Using (22) with the estimated weight vector, the approximate event-based HJB equation is

$$H(x, \hat{W}_c, \mu(\hat{x}_j), w)$$

$$= \| r^T \|^2 g_M^2(x) + \eta^2 \| m^T \|^2 W_M^2(x) + U(x, \mu(\hat{x}_j), w)$$

$$+ \hat{W}_c^T \nabla \phi(f(x) + g(x)\mu(\hat{x}_j)) + d(x)w \right) \right) \right)$$

$$= e_{cH} + (r^T (u(x) - \mu(\hat{x}_j)))^T (r^T (u(x) - \mu(\hat{x}_j)))$$

(26)

where $e_{cH}$ is a residual equation error.

Define $e_u = (r^T (u(x) - \mu(\hat{x}_j)))^T (r^T (u(x) - \mu(\hat{x}_j)))$ as the event-based transformation error. Letting the weight estimation error of the critic NN be $\hat{W}_c = W_c - \hat{W}_c$ and by combining (25) with (26), we have

$$e_{cH} = -\hat{W}_c^T \nabla \phi(f(x) + g(x)\mu(\hat{x}_j)) + d(x)w) + e_{cH} + e_u.$$

(27)

Based on experience replay [17], it is desired to choose $\hat{W}_c$ to minimize the corresponding squared residual error

$$E = \frac{1}{2} (e_{cH}^2 + \sum_{k=1}^p e^T(t_k)e(t_k))$$

where $e(t_k) = U(x(t_k), \mu(\hat{x}(t_k)), \hat{w}(t_k)) + \hat{W}_c^T(t)\sigma_k$ and $\sigma_k = \nabla \phi(x(t_k))(f(x(t_k)) + g(x(t_k))\mu(\hat{x}(t_k)) + d(x(t_k))\hat{w}(t_k))$ are stored data at time $t_k \in [t_i, t_{i+1}), i \in \mathbb{N}$, and $p$ is the number of stored samples.

1) PE-Like Condition: The recorded data matrix $F = [\sigma_1, \ldots, \sigma_p]$ contains as many linearly independent elements as the number of the critic NN’s hidden neurons, such that rank($F$) = $N$.

When using the experience replay, the PE-like condition should be satisfied, which is easily checked online. With the PE-like condition, the traditional PE condition is not required. The detailed explanation can be found in [17].

The weights of the critic NN are tuned using the standard steepest descent algorithm, which is given by

$$\hat{W}_c = -\alpha_c \nabla E \frac{\partial E}{\partial \hat{W}_c}$$

$$= -\alpha_c \sigma (\| r^T \|^2 g_M^2(x) + \eta^2 \| m^T \|^2 W_M^2(x)$$

$$+ \sigma \hat{W}_c + U(x, \mu(\hat{x}_j), \hat{w}(t)))$$

$$- \alpha_c \sum_{k=1}^p \sigma_k (\| r^T \|^2 g_M^2(x(t_k)) + \eta^2 \| m^T \|^2 W_M^2(x(t_k))$$

$$+ \sigma \hat{W}_c + U(x(t_k), \mu(\hat{x}(t_k)), \hat{w}(t_k)))$$

(28)

where $\sigma = \nabla \phi(f(x) + g(x)\mu(\hat{x}_j)) + d(x)\hat{w}(t)$, and $\alpha_c$ denotes the learning rate.

Combine (25), (27), and (28), we have

$$\dot{\hat{W}}_c = -\alpha_c \sigma (\hat{W}_c - e_{cH} - e_u)$$

$$- \alpha_c \sum_{k=1}^p \sigma_k (\hat{W}_c - e_{cH}(t_k) - e_u(t_k))$$

(29)
where $\varepsilon_c H(t_k)$ and $e_u(t_k)$ denote the residual error and the event-based transformation error at $t = t_k$, respectively.

Note that the closed-loop sampled-data system behaves as an impulsive dynamical system with the flow dynamics and jump dynamics. Define the augmented state $\Psi = [x^T, \dot{x}^T_j, \hat{W}_c^T]^T$. From (9), (10), and (29), the dynamics of the impulsive system during the flow $t \in (\tau_j, \tau_{j+1})$, $j \in \mathbb{N}$ can be described by

$$
\Psi = \begin{bmatrix}
F(\Psi) \\
0 \\
G(\Psi)
\end{bmatrix}
$$

where the nonlinear functions

$$
F(\Psi) = f(x) + g(x)\dot{\mu}(\hat{x}_j) + d(x)\sigma(x) + \frac{1}{2}g(x)\dot{\mu}(\hat{x}_j)\n + \frac{1}{2}\sigma(x)\dot{\mu}(\hat{x}_j) + d(x)\sigma(x)\n + \frac{1}{2}\sigma(x)\dot{\mu}(\hat{x}_j) + d(x)\sigma(x)
$$

$$
G(\Psi) = -a_c \sigma(\sigma^T\hat{W}_c - e_c H - e_u)
 - a_c \sum_{k=1}^p \sigma_k(\sigma_k^T\hat{W}_c - e_c H - e_u).
$$

The jump dynamics at the triggering instant $t = \tau_{j+1}$ can be given by

$$
\Psi(t) = \Psi(t^-) + \begin{bmatrix}
0 \\
-x - \hat{x}_j \\
0
\end{bmatrix}
$$

where $\Psi(t^-) = \lim_{t \to 0} \Psi(t - \varphi)$, and 0s are null vectors with appropriate dimensions.

B. Stability Analysis

In this section, the weight estimation error of the critic NN will be proved to be UUB. Meanwhile, the stability of the impulsive dynamical system based on the event-based optimal control and the augmented control will be guaranteed with a novel adaptive triggering condition. Now, we give Assumption 4, which is help to prove the stability of the closed-loop system.

Assumption 4:

1) The system dynamics $g(x)$ and $d(x)$ are upper bounded by positive constants such that $\|g(x)\| \leq \beta_{\max}$ and $\|d(x)\| \leq \beta_{\max}$.
2) The critic NN activation function and its gradient are bounded, i.e., $\|\phi(x)\| \leq \phi_{\max}$ and $\|\nabla \phi(x)\| \leq \phi_{\max}$, with $\phi_{\max}$, $\nabla \phi_{\max}$ being positive constants.
3) The gradient of critic NN activation function $\nabla \phi(x)$ is Lipschitz continuous such that $\|\nabla \phi(x) - \nabla \phi(\hat{x}_j)\| \leq L_{\phi}\|\dot{e}_j(t)\|$.
4) The critic NN ideal weight matrix is bounded by a positive constant, that is $\|\hat{W}_c\| \leq W_{\max}$.

Theorem 2: Suppose that Assumptions 1–4 hold and that the tuning law for the CT critic NN is given by (28). Then, there exists $\hat{W}_e(0)$ as discussed in Remark 10, such that the closed-loop sampled-data system (10) is asymptotically stable, and the critic weight estimation error is guaranteed to be UUB if the adaptive triggering condition is designed as

$$
\|e_j(t)\|^2 > \|R\| \frac{\|\dot{W}_e\|^2}{\ell\|\hat{W}_e\|^2} + \|R \nabla \hat{\mu}(\hat{x}_j)\|^2 - \eta_2 \|\hat{W}_e(\dot{x}_j, \hat{x}_j)\|^2 - \eta_2 \|\hat{W}_e(\dot{x}_j, \hat{x}_j)\|^2.
$$

where $\varepsilon_t \hat{W}_c(\dot{x}_j, \hat{x}_j)$ and the critic weight estimation error is guaranteed to be UUB
According to Assumptions 1 and 4, we have
\[
\|g^T(x)\nabla\phi^T(x) - g^T(\hat{x})\nabla\phi^T(\hat{x})\|^2 \\
\leq 2\|\nabla\phi(x)\left(g(x) - g(\hat{x})\right)\|^2 + 2\|g(\hat{x})(\nabla\phi(x) - \nabla\phi(\hat{x}))\|^2 \\
\leq 2(t^2_0\phi_{\text{max}}^2 + t^2_\phi\sigma_{\text{max}}^2)(e_t(t))^2
\]
and
\[
2\eta^2 w^T(x)Mw^*(x) \leq \frac{1}{2\eta^2}\|d^T(x)\nabla\phi^T(x)(W_c + \nabla c)^2 \\
\leq \frac{d^2_{\text{max}}}{2}\|d^T\nabla\phi_{\text{max}}^2(W_{\text{max}} + \nabla c_{\text{max}})^2 .
\]
Define \(\ell = t^2_0\nabla\phi_{\text{max}}^2 + t^2_\phi\sigma_{\text{max}}^2\). Then, we can obtain
\[
\hat{\pi}_{1} \leq -\ell^T Q x - \|r^T\|_2^2 g^2_\text{M}(x) - \eta^2 m^T_\text{W}^2(x) \\
+ \frac{\ell}{\|R\|^2}\|\hat{W}_c\|^2 - \|r^T \hat{\mu}(\hat{x})\|^2 + \eta^2 m^T \hat{\mu}(x))^2 \\
+ \frac{d^2_{\text{max}}}{2\eta^2}\|d^T\nabla\phi_{\text{max}}^2(W_{\text{max}} + \nabla c_{\text{max}})^2 \\
+ \frac{1}{\|R\|^2}\|\sigma_{\text{max}}^2\|\nabla\phi_{\text{max}}^2(W_{\text{max}} + \nabla c_{\text{max}})^2 .
\]
Define \(H = \sigma\sigma^T + \sum_{k=1}^p\sigma_k\sigma_k^T\). If the PE-like condition is satisfied, we have \(H > 0\). Then, we can get
\[
\hat{\pi}_{3} \leq -a_c\hat{\lambda}_{\text{min}}(H)\|\hat{W}_c\|^2 \\
+ a_c\hat{W}_c^T \left(\sigma(\epsilon_c H + \epsilon_a) + \sum_{k=1}^p\sigma_k(\epsilon_c H(t_k) + \epsilon_a(t_k))\right)
\]
Since \(g(x), \nabla\phi(x), \nabla c, \) and \(W_c\) are both bounded, we have
\(\|u(x) - \mu(\hat{x})\| = \frac{1}{2}\|R^{-1}g^T(\hat{x})\nabla\phi^T(\hat{x})W_c + \nabla c(\hat{x}) - R^{-1}g(x)(\nabla\phi^T(x)W_c + \nabla c(x))\|\) is bounded. Hence, the error \(e_u = (R^{-1}(u(x) - \mu(\hat{x})))^T(r^T(u(x) - \mu(\hat{x})))\) is also bounded. Let \(e_{\text{max}}\) be the upper bound of \(e_u\). By using Young’s inequality to the second term, (36) can be written as
\[
\hat{\pi}_{3} \leq -(a_c - 1)\hat{\lambda}_{\text{min}}(H)\|\hat{W}_c\|^2 + \frac{a_c^2}{4}\|\epsilon_c H\| + \frac{e_u}{4}\|e_{\text{max}}\|^2 .
\]
Combining (35) and (37), we can obtain
\[
\hat{L} \leq -x^T Q x - \|r^T\|_2^2 g^2_\text{M}(x) - \eta^2 m^T_\text{W}^2(x) \\
+ \frac{\ell}{\|R\|^2}\|\hat{W}_c\|^2 - \|r^T \hat{\mu}(\hat{x})\|^2 + \eta^2 m^T \hat{\mu}(x))^2 \\
+ \frac{d^2_{\text{max}}}{2\eta^2}\|d^T\nabla\phi_{\text{max}}^2(W_{\text{max}} + \nabla c_{\text{max}})^2 \\
+ \frac{1}{\|R\|^2}\|\sigma_{\text{max}}^2\|\nabla\phi_{\text{max}}^2(W_{\text{max}} + \nabla c_{\text{max}})^2 \\
+ \frac{1}{\|R\|^2}\|\sigma_{\text{max}}^2\|\nabla\phi_{\text{max}}^2(W_{\text{max}} + \nabla c_{\text{max}})^2 .
\]
If the triggering condition is designed as (32) and inequality (33) is satisfied, we can conclude \(L \leq -\beta^2\hat{\lambda}_{\text{min}}(Q)\|x\|^2 - \|r^T\|_2^2 g^2_\text{M}(x) - \eta^2 m^T_\text{W}^2(x) < 0\) under event-based control mechanism. In other words, the Lyapunov derivative is negative during the flow for \(t \in [t_{j-1}, t_j]\).

**Case II:** At the triggering instant, \(t = t_j, j \in \mathbb{N}\).

According to (34), the difference Lyapunov function candidate is written as
\[
\Delta L = L(t^-) - L(t^-) \\
= V^*(\hat{x}_j) - V^*(\hat{x}(t^-)) + V^*(\hat{x}_j) - V^*(\hat{x}_{j-1}) \\
+ \frac{1}{2}\hat{W}_c^T(\hat{x}_j)\hat{W}_c \\
- \frac{1}{2}\hat{W}_c^T(x(t^-))\hat{W}_c(x(t^-)), \ t = t_j .
\]
From (32)–(33) and (38), we can conclude that \(\dot{L} < 0\) and the system states are asymptotically stable during the flow duration. Note that the system state \(x\) is continuous for the auxiliary sampled-date system. Hence, for \(\forall t = t_j\), we have \(V^*(\hat{x}_j) \leq V^*(x(t^-))\) and \((1/2)\hat{W}_c^T(x(t^-))\hat{W}_c(x(t^-)) \leq (1/2)\hat{W}_c^T(x(t^-))\hat{W}_c(x(t^-)).\
Then, one can get
\[
\Delta L < V^*(\hat{x}_j) - V^*(\hat{x}_{j-1}) \leq -\nu((\hat{\epsilon}_j(t_j-1)))
\]
where \(\nu\) is a class-K function [35], \(\hat{\epsilon}_j(t_j-1) = \hat{x}_j - \hat{x}_{j-1}\). This implies that the Lyapunov function candidate (34) is also decreasing at triggering instants \(t = t_j, j \in \mathbb{N}\).

To sum up, if the triggering condition is defined as (32) and inequality (33) holds, we can derive the conclusion that the closed-loop impulsive system is asymptotically stable and the critic estimation error is UUB. The proof is completed.

**Remark 8:** The threshold \(\hat{\epsilon}_T\) in (32) is designed as the function of the system state vector and the critic NN weight estimates and the critic NN event-based weight estimates. That is, the triggering condition (32) is adaptive. The controller \(\hat{\mu}(\hat{x})\) is adjusted with events. Compared with the actor-critic structure in [31], a single critic network can reduce the computational burden, and the additional robust term to guarantee the closed-loop system stable is not required. In contrast to [32], the transformation error \(e_u\) is also considered for designing the weight updating law.

**Remark 9:** Note that the triggering condition (32) is utilized to approximate the optimal control policy pair \([\hat{\mu}^*(\hat{x}), \hat{w}^*(\hat{x})]^T\) for the auxiliary sampled-date system, while the triggering condition (15) in Theorem 1 is utilized to guarantee the robust stabilization of the original uncertain system with the obtained optimal control policy \(\hat{\mu}^*(\hat{x})\).

**Corollary 1:** The approximate optimal control policy pair \([\hat{\mu}^*(\hat{x}), \hat{w}^*(\hat{x})]^T\) in (23) and (24) converges to the optimal control inputs \([\mu^*(\hat{x}), w^*(\hat{x})]^T\) with a finite bound.

**Proof:** From (11), (21), and (24), we have
\[
\hat{\mu}(\hat{x}_j) - \mu^*(\hat{x}_j) \leq \frac{1}{2}R^{-1}g^T(\hat{x}_j)\nabla\phi^T(\hat{x}_j)\hat{W}_c \\
+ \frac{1}{2}R^{-1}g^T(\hat{x}_j)\nabla\phi^T(\hat{x}_j)\nabla c(\hat{x}_j). (39)
\]
According to Assumption 4 and Theorem 2, we know that the first and second terms in (39) are all bounded. That is
\[
\|\hat{\mu}(\hat{x}_j) - \mu^*(\hat{x}_j)\| \leq \frac{1}{2}\|R\|\|\nabla\phi_{\text{max}}(\Pi_{\text{max}} + \nabla c_{\text{max}})\| \leq \epsilon_{\mu}.
\]
where \(\Pi_{\text{M}}\) is defined in (33), and \(\epsilon_{\mu}\) is the finite bound. Similarly, we can obtain the finite bound.
This approximate optimal control policy is achieved with the converged critic NN for the auxiliary sampled-data system is shown in Fig. 1, where an approximate optimal control policy pair \((\hat{\mu}^*(\hat{x}_j), \hat{w}^*(x))\) is obtained by (32) is low bounded by

\[
\delta \tau_{\text{min}} \geq \frac{P}{Y (P + 1)} > 0
\]

where

\[ P = \frac{((4\eta^2(1 - \beta^2)) M ||\lambda_{\text{min}}(Q) - \ell_2^2 \| \hat{W}_c \|^2)/(4\eta^2 \| M \|))^{1/2}}{} \]

and

\[ Y = \ell_f + \frac{((\ell_2^2)/(2\eta^2 \| M \|)) + ((d_{\text{max}})/(2\eta^2 \| M \|)) \ell_\phi \| \hat{W}_c \|}{\ell_\phi}. \]

Proof: According to the ETC mechanism, at the triggering instant, we have \(\|e_j(t_j+1)\| = \|\hat{x}_j(t, \hat{x}_j, \hat{W}_c)\|\). Combined with (32), we can get

\[
\ell \| \hat{W}_c \|^2 \| e_j(t) \|^2 + \eta^2 m^T \hat{w}(x) \| = (1 - \beta^2) \lambda_{\text{min}}(Q) \| x \|^2 + \| x^T \hat{\mu}(\hat{x}_j) \|^2.
\]

Based on (23) and Assumption 4, we have

\[
\ell \| \hat{W}_c \|^2 \| e_j(t) \|^2 + \frac{\ell_\phi^2 \| \hat{W}_c \|^2}{4 \eta^2 \| M \|} \| x \|^2
\]

\[
\geq (1 - \beta^2) \lambda_{\text{min}}(Q) \| x \|^2 + \| x^T \hat{\mu}(\hat{x}_j) \|^2.
\]

Then, we can obtain

\[
\ell \| \hat{W}_c \|^2 \| e_j(t) \|^2 \geq \left(1 - \beta^2\right) \lambda_{\text{min}}(Q) - \frac{\ell_\phi^2 \| \hat{W}_c \|^2}{4 \eta^2 \| M \|} \| x \|^2.
\]

Since at the \(j\)th triggering instant \(e_j(t) = 0\), so the time of \(((\|e_j(t)\|)/(\|x\|))\) growing from 0 to \(P\) provides a lower bound for the minimum interevent time.

Based on Assumption 4, we have

\[
\| \hat{\mu}(\hat{x}_j) \| \leq \frac{d_{\text{max}}}{2 \| R \|} \ell_\phi \| \hat{W}_c \| \| \hat{x}_j \|
\]

\[
\| \hat{w}(x) \| \leq \frac{d_{\text{max}}}{2 \eta^2 \| M \|} \ell_\phi \| \hat{W}_c \| \| x \|.
\]

With (40), the system state satisfies

\[
\| \hat{x} \| = \| f(x) + g(x) \hat{\mu}(\hat{x}_j) + d(x) \hat{w}(x) \|
\]

\[
\leq \ell_f \| x \| + \frac{\ell_\phi}{2 \| R \|} \| \hat{W}_c \| \| x \| + \ell_f \| x \|
\]

\[
+ \frac{d_{\text{max}}^2}{2 \eta^2 \| M \|} \ell_\phi \| \hat{W}_c \| \| x \|
\]

\[
\leq \Upsilon(\| x \| + \| e_j \|).
\]

Remark 10: Note that the initialization of the critic NN weights \(\hat{W}_c\) plays an important role in approaching the stabilizing solution of the HJB equation. For the time-based online ADP algorithm, the weights of critic network can be initialized to zero by introducing a rigorous assumption in [13]. However, for the event-based online ADP algorithm, how to choose the initial weights of critic network is still an interesting and practical topic. Similar to [36], the initial weights of critic network are chosen experimentally in this paper.

C. Low Bound on Intersample Times

For the CT systems with event-based controller, the analysis on Zeno behavior (an infinite number of discrete transitions occur in a finite time interval) is needed. Then, we give Theorem 3 to guarantee that the minimal intersample time \(\delta \tau_{\text{min}} = \min_{j \in \mathbb{N}}(t_{j+1} - t_j)\) is bounded by a nonzero positive constant. In other words, the Zeno behavior is excluded during the learning process.

**Theorem 3:** Consider the auxiliary sampled-data system (10) with the triggering condition (32), the minimal intersample time \(\delta \tau_{\text{min}}\) obtained by (32) is low bounded by

\[
\delta \tau_{\text{min}} \geq \frac{P}{Y (P + 1)} > 0
\]

where

\[
P = \frac{((4\eta^2(1 - \beta^2)) M ||\lambda_{\text{min}}(Q) - \ell_2^2 \| \hat{W}_c \|^2)/(4\eta^2 \| M \|))^{1/2}}{} \]

and

\[
Y = \ell_f + \frac{((\ell_2^2)/(2\eta^2 \| M \|)) \ell_\phi \| \hat{W}_c \|}{\ell_\phi}.
\]

Proof: According to the ETC mechanism, at the triggering instant, we have \(\|e_j(t_j+1)\| = \|\hat{x}_j(t, \hat{x}_j, \hat{W}_c)\|\). Combined with (32), we can get

\[
\ell \| \hat{W}_c \|^2 \| e_j(t) \|^2 + \eta^2 m^T \hat{w}(x) \|^2
\]

\[
= (1 - \beta^2) \lambda_{\text{min}}(Q) \| x \|^2 + \| x^T \hat{\mu}(\hat{x}_j) \|^2.
\]

Based on (23) and Assumption 4, we have

\[
\ell \| \hat{W}_c \|^2 \| e_j(t) \|^2 + \frac{\ell_\phi^2 \| \hat{W}_c \|^2}{4 \eta^2 \| M \|} \| x \|^2
\]

\[
\geq (1 - \beta^2) \lambda_{\text{min}}(Q) \| x \|^2 + \| x^T \hat{\mu}(\hat{x}_j) \|^2.
\]

Then, we can obtain

\[
\ell \| \hat{W}_c \|^2 \| e_j(t) \|^2 \geq \left(1 - \beta^2\right) \lambda_{\text{min}}(Q) - \frac{\ell_\phi^2 \| \hat{W}_c \|^2}{4 \eta^2 \| M \|} \| x \|^2.
\]

Since at the \(j\)th triggering instant \(e_j(t) = 0\), so the time of \(((\|e_j(t)\|)/(\|x\|))\) growing from 0 to \(P\) provides a lower bound for the minimum interevent time.
Based on the UUB conclusion in Theorem 2, the norm of critic estimation $\| \hat{W}_c \|$ is bounded. So does $\Upsilon$. According to [37], the dynamics of $((\| e_j(t) \|)/(\| x \|))$ satisfies

$$\frac{d}{dt} \| e_j(t) \| \leq \left( 1 + \frac{e_j(t)}{\| x \|} \right) \| x \|.$$ 

Combined with (41), we can obtain

$$\frac{d}{dt} \| e_j(t) \| \leq \Upsilon \left( 1 + \frac{e_j(t)}{\| x \|} \right)^2.$$ 

From [37, Th. III.1], the time that $((\| e_j(t) \|)/(\| x \|))$ evolves from 0 to $P$ is greater than $(P/(\Upsilon(P + 1)))$. The proof is complete.

Remark 11: Note that $P$ can be positive by selecting appropriate matrices $Q$ and $M$. From Theorem 3, we know that Zeno behavior is excluded during the learning process. However, some unnecessary events are usually triggered when the system states are fluctuated inside a small bound near the equilibrium point. According to [34], a dead-zone operator can be used to solve this problem effectively.

V. SIMULATION

To demonstrate the effectiveness of the developed algorithm, we choose two examples for numerical experiments.

Example 1: The first example is considered as follows [7]:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u + \begin{bmatrix} 0.2 \\ 0 \end{bmatrix} W(x)$$

where $W(x) = \lambda_1 x_1 \cos(1/(x_2 + \lambda_2)) + \lambda_2 x_2 \sin(\lambda_4 x_1 x_2)$, and $\lambda_1, \lambda_2, \lambda_3, \text{and} \lambda_4$ are the unknown parameters with $\lambda_1 \in [-1, 1]$, $\lambda_2 \in [-100, 100]$, $\lambda_3 \in [-0.2, 1]$, and $\lambda_4 \in [-100, 0]$. The last term reflects the unmatched uncertainty in the system.

Clearly

$$g^+(x) = (g^T(x)g(x))^{-1}g^T(x)$$

$$= g^T(x) = [0, 1]$$

$$(I - g(x)g^+(x))k(x) = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0.2 \\ 0 \end{bmatrix} = \begin{bmatrix} 0.2 \\ 0 \end{bmatrix}$$

$$\| W(x) \|^2 \leq x_1^2 + x_2^2 \leq W_M^2(x)$$

and

$$2\| g^+(x)k(x)W(x) \|^2 = 0 \leq g_M^2(x).$$

Let $Q, R, r, \text{and} m$ be the identity matrices with appropriate dimensions. We experimentally choose $\eta = 1, \rho = 10, \beta = 0.1, \ell = 4, \text{and} L = 3$. Then, the corresponding optimal control problem is as follows. For the auxiliary sampled-data system

$$\begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} \mu(\hat{x}_j) + \begin{bmatrix} 0.2 \\ 0 \end{bmatrix} w(x)$$

find out the optimal control laws $[\mu^T(\hat{x}_j), w^T(x)]^T$ to minimize the cost function

$$\int_0^\infty [2x^Tw + u^T(x)u(x) + w^T w]dt.$$
Based on the converged weights \( \hat{W}_c \), we can obtain the near-optimal control laws as

\[
\begin{bmatrix}
\hat{\mu}^\ast(x_j) \\
\hat{w}^\ast(x)
\end{bmatrix} = 
\begin{bmatrix}
-1/2 \begin{bmatrix} 0 & 1 \end{bmatrix} \nabla \phi^T (x_j) \hat{W}_c \\
-1/2 \begin{bmatrix} 0.2 & 0 \end{bmatrix} \nabla \phi^T (x) \hat{W}_c
\end{bmatrix}.
\] (42)

From [7], the optimal control laws are given as

\[
\begin{bmatrix}
u^\ast(x) \\
w^\ast(x)
\end{bmatrix} = 
\begin{bmatrix}
-1.2906x_1 - 2.1247x_2 \\
-0.5783x_1 - 0.2581x_2
\end{bmatrix}.
\] (43)

Now, we apply the near-optimal control laws (42) with the triggering condition (15) and the optimal control laws (43) for the uncertain nonlinear system. Set the initial state be \( x_0 = [1, -1]^T \), and the sampling time be 0.05s. The simulation results for the following four cases are given in Figs. 5–8.

Case 1: \( \lambda_1 = -1, \lambda_2 = -100, \lambda_3 = 0, \) and \( \lambda_4 = -100 \).
Case 2: \( \lambda_1 = 0.2, \lambda_2 = 100, \lambda_3 = 1, \) and \( \lambda_4 = -1 \).
Case 3: \( \lambda_1 = 0, \lambda_2 = 0, \lambda_3 = 0, \) and \( \lambda_4 = 0 \).
Case 4: \( \lambda_1 = -0.2, \lambda_2 = -100, \lambda_3 = -0.2, \) and \( \lambda_4 = -100 \).

From Figs. 5–8, we can observe that the state trajectories are different from the above four disparate perturbation parameters \( (\lambda_1, \lambda_2, \lambda_3, \lambda_4) \). However, they can both converge to the equilibrium point under the near-optimal control policies (42) and the triggering condition (15). Furthermore, the near-optimal control input is almost a discretized version of the optimal control input for these cases. Evidently, the near-optimal controller is robust and adjusted with events.
Example 1. After a sufficient learning process, the parameters matrices with appropriate dimensions. We choose $\eta$ structure 3-21-1. The critic NN activation function is chosen as

$$\phi(x) = [x_1^2 x_2^2 x_3 x_1 x_2 x_3 x_2 x_3 x_3 x_2 x_3 x_1 x_2 x_3] \ ,$$

where $W(x) = \lambda_5 x_1 \sin x_2 \cos x_3$ denotes the system uncertainty, and $\lambda_5$ is the unknown parameter with $\lambda_5 \in [-1, 1]$.

Clearly

$$g^+(x) = (g^T(x)g(x))^{-1}g^T(x) = g^T(x) = [0, 1, 0]^T$$

$$k(x) = [0, 0, 1]^T$$

and

$$\|W(x)\|^2 \leq x_1^2 = W_M^2(x)$$

and

$$2\|g^+(x)k(x)W(x)\|^2 = 0 \leq g_M^2(x).$$

In this example, $Q = 2I$, $R$, $r$, and $m$ are still the identity matrices with appropriate dimensions. We choose $\eta = 1$, $p = 30$, $\beta = 0.1$, $\ell = 4$, and $L = 5$. For the auxiliary sampled-data system, we choose the critic NN with structure 3-21-1. The critic NN activation function is chosen as

$$\phi(x) = [x_1^2 x_2^2 x_3 x_1 x_2 x_3 x_2 x_3 x_3 x_2 x_3 x_1 x_2 x_3 x_1 x_2 x_3] \ .$$

Let the initial state be $x_0 = [1, -1, 0.5]^T$. The learning rate and sampling time are chosen the same as the first case of Example 1. After a sufficient learning process, the parameters of critic NN converge to

$$\hat{W}_c = [0.4821 \ 0.6176 \ 0.3012 \ 0.4669 \ 0.1583 \ 0.1312 \ 0.0869 \ 0.0125 \ 0.1816 \ 0.1638 \ 0.1047 \ 0.0162 \ -0.1128 \ 0.0147 \ -0.0296 \ -0.1458 \ 0.0765 \ -0.0237 \ 0.1358 \ 0.0343]^T \ .$$

The evolution of triggering condition (32) is shown in Fig. 9. The sampling period for the control input is shown in Fig. 10 with the minimal intersample time 0.25 s. In particular, the event-based controller uses 32 samples, while the time-triggered controller uses 1000 samples.

Let the scalar parameter $\lambda_5 = -1$. Similarly, we apply the obtained near-optimal control laws with the triggering condition (15) based on the converged critic NN weights and the optimal control law provided in [13] for the uncertain system (44), respectively. Choose the same initial state and sampling period at the learning process. The simulation results are given in Fig. 11. One can see that the near-optimal control input is a discretized version of the optimal control input.

Comparing the event-based ADP and the traditional ADP in [13], the total number of computations in terms of addition and multiplications [30] at the controller for the same

### Table I

**COMPARISONS OF COMPUTATIONAL LOAD BETWEEN TRADITIONAL ADP AND EVENT-BASED ADP**

<table>
<thead>
<tr>
<th>Method</th>
<th>Traditional ADP in [13]</th>
<th>Event-based ADP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sampling instants</td>
<td>1000</td>
<td>32</td>
</tr>
<tr>
<td>Number of additions and multiplications at every sampling instant for control input</td>
<td>8</td>
<td>23</td>
</tr>
<tr>
<td>Total number of computation</td>
<td>8000</td>
<td>992</td>
</tr>
</tbody>
</table>
simulation time is shown in Table I. Evidently, the computational burden is reduced using the event-based ADP. These simulation results verify the effectiveness of the developed control scheme.

VI. CONCLUSION AND DISCUSSION
In this paper, we propose an event-based ADP algorithm for the robust control of uncertain nonlinear systems. The robust control problem is described as an optimal control of an auxiliary system. For implementation purpose, a critic NN is constructed to approximate the optimal value function. Then, the stability analysis of the closed-loop system is given. The Zeno behavior is excluded during the learning phase.

Our future theoretic work is to avoid the continuous monitoring of the triggering condition for uncertain nonlinear systems and consider relaxing the initial stabilizing control policy in the event-based online algorithm. For the potential practical application, we hope to further investigate the cooperative ACC system combining the proposed event-based ADP algorithm and our work about the ACC system in [18].

REFERENCES