Adaptive resilient event-triggered control design of autonomous vehicles with an iterative single critic learning framework

Zhang, Kun; Su, Rong; Zhang, Huaguang; Tian, Yunlin

2021

Zhang, K., Su, R., Zhang, H. & Tian, Y. (2021). Adaptive resilient event-triggered control design of autonomous vehicles with an iterative single critic learning framework. IEEE Transactions On Neural Networks and Learning Systems, 32(12), 5502-5511.
<https://dx.doi.org/10.1109/TNNLS.2021.3053269>

<https://hdl.handle.net/10356/152244>

<https://doi.org/10.1109/TNNLS.2021.3053269>

© 2021 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works. The published version is available at:
<https://doi.org/10.1109/TNNLS.2021.3053269>.

Downloaded on 26 Apr 2023 17:17:15 SGT

Adaptive Resilient Event-Triggered Control Design of Autonomous Vehicles with an Iterative Single Critic Learning Framework

Kun Zhang, *Member, IEEE*, Rong Su, *Senior Member, IEEE*, Huaguang Zhang, *Fellow, IEEE*, and Yunlin Tian

Abstract—This paper investigates the adaptive resilient event-triggered control for rear wheel drive autonomous (RWDA) vehicles based on an iterative single critic learning framework, which can effectively balance the frequency/changes in adjusting the vehicle's control during the running process. According to the kinematic equation of RWDA vehicles and desired trajectory, the tracking error system during autonomous driving process is firstly built, where the denial-of-service (DoS) attacking signals are injected in the networked communication and transmission. Combining the event-triggered sampling mechanism and iterative single critic learning framework, a new event-triggered condition is developed for the adaptive resilient control algorithm, and the novel utility function design are considered for driving the autonomous vehicle, where the control input can be guaranteed into an applicable saturated bound. Finally, we apply the new adaptive resilient control scheme to a case study of driving the RWDA vehicles, and the simulation results illustrate the effectiveness and practicality successfully.

Index Terms—Autonomous vehicle, optimal control, resilient control, adaptive dynamic programming (ADP), event-triggered control, tracking control

I. INTRODUCTION

WITH the rapid development of internet of things, artificial intelligence (AI) and new energy technologies, a new era of automobile intelligence has come quietly. After the “Alpha-Go” algorithm defeated the champion of human beings in man-machine games, the AI technique, which is based on reinforcement learning (RL) or adaptive dynamic programming (ADP) methods, has recently drawn much attention of engineers and scholars from the realm of science and technology [1]–[4]. That is because that RL and ADP methods have the capability to solve the optimal control, which can maximize or minimize the cost function while stabilizing the

system dynamics. Partially due to the feasibility of solving optimal control problems by RL and ADP methods, which have found good applications in autonomous vehicle driving [5], [6]. However, it is still in the infancy stage, and many pivotal control issues on the autonomous vehicles are still far from perfect.

Generally, the energy expenditure in the driving process plays a key role in determining the main performance of an autonomous vehicle. One way is the optimal control method, which gives the maximum or minimum of the performance index. Based on the principle of optimality, the optimal control satisfies a mathematical partial differential equation [7]–[10], which is difficult to be solved by an analytical method directly. To overcome this challenge, many scholars try to find an approximated solution by iterative learning methods [11]–[13], which depend on RL and ADP architecture. In particular, for the traditional optimal tracking control problem with asymptotically stable trajectories, the control input must be divided into steady-state part and feedback part to guarantee the finiteness of the cost function [14]–[16]. To better measure the control input, the discounted cost function was designed by some researchers [17]–[19] for an augmented tracking system reconstructed by combining system and reference dynamics. Another way is the event-triggering control method, which can reduce the cost of control operation in the driving process. Distinct from the time-driven methods, the event-triggering mechanism can optimize the control cost by decreasing data sampling frequency. With this characteristic, many event-driven structures and methods were developed [20]–[22], and [23] discussed a self-learning optimal regulation, where an event-based adaptive critic algorithm was developed with an event-driven formulation. However, the optimal cost functions were not further analyzed according to the event sampling processes at this stage.

It should be mentioned that the communication networks in an autonomous vehicle system, which may suffer denial-of-service (DoS) attacks, affect the driving path and feedback control policy by transmitting the driving information. To overcome the DoS attacks in networked cyber-physical systems, some resilient control methods were developed to resist their negative effects [24], [25]. Among the complex systems in cyber-environment, resilient distributed control schemes were developed for multi-agent systems [26]–[28], and [29] researched the formation control problem of nonlinear multi-agent systems under DoS attacks and achieved the contribution with hybrid event-triggering strategies. Considering

This work was supported by the National Postdoctoral Program for Innovative Talents of China (BX20200357), the China Postdoctoral Science Foundation under Grant 2020M680718, the Singapore National Research Foundation Delta-NTU Corporate Lab Program (DELTA-NTU CORP-SMARP2), and Singapore Ministry of Education Tier 1 Academic Research Grant 2013-T1-002-177.

K. Zhang is with the Key Laboratory of Systems and Control, Institute of Systems Science, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100049, China (e-mail: nukgnahz@163.com).

R. Su is with the the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore 639798 (rsu@ntu.edu.sg).

H. Zhang is with the State Key Laboratory of Synthetical Automation for Process Industries, Northeastern University, and also with the School of Information Science and Engineering, Northeastern University, Shenyang 110819, China (e-mail: hgzhang@ieee.org).

Y. Tian is with the Faculty of Business, University of Wollongong, NSW 2522, Australia (e-mail: yt668@uowmail.edu.au).

the communication protocol of networked systems, the state-saturated resilient strategies were investigated in [30], [31] to against the deception attacks, and [32] designed a switching-dependent controller for the asynchronously switched systems with adaption to fast-switching perturbation. The resilient stabilization of switched discrete-time systems was studied in [33] against adversarial switching, and [34] proposed a class of new convex reliable stabilization conditions for the influence of sensor faults. Besides, according to the optimal control and game theory, some resilient methods were developed to mitigate impacts of the attacks and perturbations in [35]–[37], and [38] designed an event-triggering mechanism to make a tradeoff between the transmission efficiency and tolerable attack intensity, which indicates the potential advantages of event-triggering resilient control methods.

To solve the control scheme with less energy expenditure in the vehicle's driving process, and motivated by the ADP techniques, a new resilient event-triggered control scheme for the autonomous driving system of vehicles is proposed by using the iteration single critic learning framework. The main importance and contributions in this paper can be summarized as follows: (1) A novel autonomous driving system is structured with considering the DoS attacks, and by combining event-triggered sampling mechanism and iterative single critic learning framework, the driving control process involves less energy expenditure. (2) Different from some existing event-triggering mechanisms in ADP methods, such as [20]–[23], the control scheme is developed with a specific sampling mechanism, where the cost function under event-triggered processes is analyzed by sampling intervals, which is the first of their kinds. (3) The adaptive resilient event-triggered control algorithm is first developed for the autonomous vehicles, which effectively balance the frequency/changes in adjusting the vehicle's control during the running process.

The rest of this paper is organized as follows. The main design is inspired by how to drive the autonomous vehicle running with less energy expenditure, as presented in Section I, where event-triggering mechanism and resilient control problem are considered. The system model of a RWDA vehicle is built based on the kinematic equation in Section II, where DoS attacks occur in the dynamic system. Section III gives the event-triggered sampling mechanism, and derives the cost function and control policy under the event-triggering processes. A new adaptive tracking control algorithm is proposed in Section IV for the autonomous driving system of vehicles, and the single critic learning framework is developed. The novel adaptive resilient event-triggered control algorithm is implemented in a simulated case study shown in Section V, and conclusions are drawn in Section VI.

II. PROBLEM FORMULATION AND PRELIMINARIES

A. The Dynamic Modeling of an Autonomous Vehicle System

In this paper, we consider the typical rear wheel drive autonomous (RWDA) vehicle [39], and the normal kinematic

equation of the RWDA vehicle is

$$\begin{cases} \dot{x}(t) = v_x(t) \cos(\theta(t)) - d_r w_x(t) \sin(\theta(t)) \\ \dot{y}(t) = v_x(t) \sin(\theta(t)) + d_r w_x(t) \cos(\theta(t)) \\ \dot{\theta}(t) = w_x(t) \end{cases} \quad (1)$$

where $x(t)$ is the horizontal position of the mass center of the vehicle in the inertial reference frame, $y(t)$ is the vehicle's vertical position, $\theta(t)$ is its orientation, $v_x(t)$ represents the longitudinal velocity of the mass center in the body fixed frame, $w_x(t)$ denotes the yaw angular velocity (around Z axis, and perpendicular to the $X - Y$ plane), and d_r is the direct distance from the mass center to the rear axle in the vehicle.

To make an RWDA vehicle follow a desired trajectory, we assume that the desired reference trajectory is generated by fixed longitudinal velocity $v_r(t)$ and yaw angular velocity $w_r(t)$ of the vehicle, and the reference dynamics can be presented as:

$$\begin{cases} \dot{x}_r(t) = v_r(t) \cos(\theta_r(t)) - d_r w_r(t) \sin(\theta_r(t)) \\ \dot{y}_r(t) = v_r(t) \sin(\theta_r(t)) + d_r w_r(t) \cos(\theta_r(t)) \\ \dot{\theta}_r(t) = w_r(t) \end{cases} \quad (2)$$

where $x_r(t)$, $y_r(t)$ and $\theta_r(t)$ indicate the desired horizontal position, vertical position and the vehicle's orientation, respectively.

In the local coordinate system of a vehicle, we define that the horizontal position error is $x_e(t)$, vertical position error is $y_e(t)$, and heading direction error is $\theta_e(t)$. Then, the tracking error vector of the vehicle during the autonomous driving becomes:

$$\begin{bmatrix} x_e(t) \\ y_e(t) \\ \theta_e(t) \end{bmatrix} = \begin{bmatrix} \cos(\theta(t)) & \sin(\theta(t)) & 0 \\ -\sin(\theta(t)) & \cos(\theta(t)) & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_r(t) - x(t) \\ y_r(t) - y(t) \\ \theta_r(t) - \theta(t) \end{bmatrix}. \quad (3)$$

The RWDA vehicle's free-body diagram and tracking trajectory are presented in Fig. 1, where the reaction forces on each individual wheel are summed up at their mid-axes and the reference path is given.

Fig. 1. The RWDA Vehicle's Free-body Diagram and Tracking Trajectory

The control objective of this paper is to find a pair of control inputs $v_x(t)$ and $w_x(t)$ of RWDA vehicles, which stabilize the tracking errors and resist the attacking signals from the vehicle's communication system. Generally, there exist a lot

of DoS attacks in networked systems, and the autonomous vehicle is particularly vulnerable to sensor and actuator attacks as it heavily rely on proper data communication. The tracking error dynamics with uncertain attacking signals are presented in the follows.

B. The Resilient Control Problem in the Autonomous Driving

According to the RWDA vehicle's free body diagram and tracking error equation (3), we compute the tracking error dynamics of a RWDA vehicle.

Theorem 1. *For a nominal RWDA vehicle system, the tracking error dynamic can be derived as:*

$$\dot{s}(t) = f(s(t)) + g(s(t))u(t) \quad (4)$$

where $s(t) = [x_e(t), y_e(t), \theta_e(t)]^T$, $u(t) = \begin{bmatrix} w_x(t) \\ v_x(t) \end{bmatrix}$,
 $f(t) = \begin{bmatrix} v_r(t) \cos(\theta_e(t)) \\ v_r(t) \sin(\theta_e(t)) \\ w_r(t) \end{bmatrix}$, and $g(t) = \begin{bmatrix} y_e(t) & -1 \\ -x_e(t) - d_r & 0 \\ -1 & 0 \end{bmatrix}$.

Proof. Firstly, based on the tracking error equation (3), using the vehicle's dynamic (1) and reference dynamic (2), the error dynamic $\dot{x}_e$ yields:

$$\begin{aligned} \dot{x}_e &= \cos(\theta)[\dot{x}_r - \dot{x}] + \sin(\theta)[\dot{y}_r - \dot{y}] \\ &\quad - \sin(\theta)\dot{\theta}[x_r - x] + \cos(\theta)\dot{\theta}[y_r - y] \\ &= \cos(\theta_r - \theta_e)\dot{x}_r - \cos(\theta)[v_x \cos(\theta) - d_r w_x \sin(\theta)] \\ &\quad + \sin(\theta_r - \theta_e)\dot{y}_r - \sin(\theta)[v_x \sin(\theta) + d_r w_x \cos(\theta)] \\ &\quad - \sin(\theta)w_x[x_r - x] + \cos(\theta)w_x[y_r - y] \\ &= \dot{x}_r[\cos(\theta_r)\cos(\theta_e) + \sin(\theta_r)\sin(\theta_e)] + y_e w_x \\ &\quad + \dot{y}_r[\sin(\theta_r)\cos(\theta_e) - \cos(\theta_r)\sin(\theta_e)] - v_x. \end{aligned} \quad (5)$$

From the RWDA vehicle's free body diagram and modeling process, one has $\dot{x}_r \sin(\theta_r) = \dot{y}_r \cos(\theta_r)$, thus, it becomes

$$\begin{aligned} \dot{x}_e &= y_e w_x + [v_r \cos(\theta_r) - d_r w_r \sin(\theta_r)] \cos(\theta_r) \cos(\theta_e) \\ &\quad - v_x + [v_r \sin(\theta_r) + d_r w_r \cos(\theta_r)] \sin(\theta_r) \cos(\theta_e) \\ &= y_e w_x - v_x + v_r \cos(\theta_e). \end{aligned} \quad (6)$$

Secondly, for the tracking error y_e , there is

$$\begin{aligned} \dot{y}_e &= -\sin(\theta)[\dot{x}_r - \dot{x}] + \cos(\theta)[\dot{y}_r - \dot{y}] \\ &\quad - \cos(\theta)\dot{\theta}[x_r - x] - \sin(\theta)\dot{\theta}[y_r - y] \\ &= -\sin(\theta_r - \theta_e)\dot{x}_r + \sin(\theta)[v_x \cos(\theta) - d_r w_x \sin(\theta)] \\ &\quad + \cos(\theta_r - \theta_e)\dot{y}_r - \cos(\theta)[v_x \sin(\theta) + d_r w_x \cos(\theta)] \\ &\quad - \cos(\theta)w_x[x_r - x] - \sin(\theta)w_x[y_r - y] \\ &= -d_r w_x - \dot{x}_r[\sin(\theta_r)\cos(\theta_e) - \cos(\theta_r)\sin(\theta_e)] \\ &\quad - x_e w_x + \dot{y}_r[\cos(\theta_r)\cos(\theta_e) + \sin(\theta_r)\sin(\theta_e)] \\ &= -d_r w_x - x_e w_x + v_r \sin(\theta_e). \end{aligned} \quad (7)$$

Finally, based on the dynamics (1) and (2), the tracking error dynamic $\dot{\theta}_e(t)$ can be directly derived by

$$\dot{\theta}_e(t) = \dot{\theta}_r(t) - \dot{\theta}(t) = w_r(t) - w_x(t). \quad (8)$$

Thus, we have

$$\begin{bmatrix} \dot{x}_e(t) \\ \dot{y}_e(t) \\ \dot{\theta}_e(t) \end{bmatrix} = \begin{bmatrix} v_r(t) \cos(\theta_e(t)) + y_e(t)w_x(t) - v_x(t) \\ v_r(t) \sin(\theta_e(t)) - x_e(t)w_x(t) - d_r w_x(t) \\ w_r(t) - w_x(t) \end{bmatrix}, \quad (9)$$

which gives the tracking error dynamic (4), and completes the proof. $\square$

In the autonomous driving and operation process, the tracking error system (4) of RWDA vehicles generally suffers uncertain attacking signals, where the DoS attacks are injected into the tracking dynamic via communication networks as the signal flow shown in Fig. 2.

Fig. 2. The Autonomous Driving System in Suffering of Uncertain DoS Attacking Signals

Then, the general autonomous driving system of RWDA vehicles can be rewritten as

$$\dot{s}(t) = f(s(t)) + \alpha(t) + g(s(t))u(t) \quad (10)$$

where $\|g(s(t))\| \neq 0$ for all $s(t)$, $u(t) = \mu(t) + u_r(t) + \beta(t)$ is the actual system input working on the vehicles, $\mu(t)$ indicates the resilient tracking control policy, $u_r(t) = [w_r(t), v_r(t)]^T$ is the reference policy, $\alpha(t) = [\alpha_x(t), \alpha_y(t), \alpha_\theta(t)]^T$ and $\beta(t) = [\beta_1(t), \beta_2(t)]^T$ denote the uncertain sensor and actuator attacking signals in the tracking dynamic functions, respectively.

In the tracking driving process, the reference policy $u_r(t)$ is used to maintain the tracking trajectory, which is a given reference in the vehicles, and the resilient tracking control policy $\mu(t)$ is used to decrease the tracking errors and overcome the uncertain signals.

Assumption 1. *In this paper, the sensor and actuator attacking signals are assumed to be state dependent, which can be parameterized as $\alpha(t) = \varpi(t)s(t)$ and $\beta(t) = W(t)\varphi(s(t))$ for all $t \geq 0$. Similar to [41], $\varpi(t)$ and $W(t)$ are unknown time-varying matrices satisfying $\|\varpi(t)\|_2 \leq b_1$, $\|W(t)\|_2 \leq b_2$ with constants $b_1, b_2 > 0$, and $\varphi(\cdot)$ is an unknown and bounded nonlinear function.*

To stabilize the tracking error dynamic (10) under the DoS attacks, we design the resilient control scheme for RWDA vehicles by using the iterative single critic learning framework, which is developed in the following section.

III. THE EVENT-TRIGGERED RESILIENT CONTROL AND STABILITY ANALYSIS OF SAMPLING MECHANISM DESIGN

A. The Event-Triggered Resilient Optimal Control with Saturating Bound

For the tracking error dynamics of RWDA vehicles, the uncertain attacking signals are considered in the communication networks. By this way, we further rewrite the autonomous driving system (10) as:

$$\begin{aligned}\dot{s}(t) &= f(s(t)) + g(s(t))(\mu(t) + u_r(t)) + \sigma(t) \\ &= \bar{f}(s(t)) + g(s(t))\mu(t) + \sigma(t)\end{aligned}\quad (11)$$

where $\bar{f}(s(t)) = f(s(t)) + g(s(t))u_r(t)$ is the desired dynamic part, $\sigma(t) = \alpha(t) + g(t)\beta(t)$ is the overall attacking signal. Based on the aforementioned assumption, it can be obtained that $\|\sigma(t)\| \leq \gamma\|g(t)\|$ with a constant $\gamma > 1$. As we have pointed out above, the reference policy $u_r(t)$ is given according to the desired reference, which is generally set in the driving system as a desired reference. Besides, in a conventional tracking control problem, the resilient tracking control policy $\mu(t)$ is designed as a time-driven feedback control policy, and here we will give a new event-triggered tracking control scheme.

Considering the process of information transmission in communication networks, the event-triggered sampling mechanism is developed with a monotonically increasing time sequence as $\{z_i\}_{i=0}^{+\infty} = \{z_0, z_1, z_2, \dots, z_i, \dots\}$, $i \in \mathcal{N}$, where $z_0 = 0$ and the system state in the control policy is updated at each triggering instant. By this way, an event-triggered feedback control is designed as $\mu(t, z_i) = \mu(t, s(z_i))$ for all the time $t \in [z_i, z_{i+1})$, where $s(z_i)$ is the state at time instant z_i , and the zero-order hold can keep the control in continuous at every time instant z_i .

Let the system state $s(t)$ be the event-triggered state, thus, based on the event-triggering mechanism, we define the event-triggered condition as follows:

$$\mathcal{F}(s(t)) \leq s_T(t, z_i) \quad (12)$$

where z_i is the latest triggering instant at time t , and $s_T(t, z_i)$ is the triggering threshold. And the event-triggered state satisfies

$$s(t) = s(z_i) - e(t) \quad (13)$$

where $e(t) = \begin{cases} 0, & t = z_i, i \in \mathcal{N} \\ s(z_i) - s(t), & t \in (z_i, z_{i+1}) \end{cases}$ denotes the triggering state error. Then, the event-triggered control policy becomes $\mu(t, z_i) = \mu(s(t) + e(t))$.

Define the infinite horizon integral performance index for the system dynamic (11) by

$$\begin{aligned}J(s(0)) &= \\ &\sum_{i=0}^{+\infty} \int_{z_i}^{z_{i+1}} [\gamma^{-\varrho\tau} \|g(\tau)\|^2 + \mathcal{Q}(s(\tau)) + U(\mu(\tau, z_i))] d\tau\end{aligned}\quad (14)$$

where $\varrho > 0$ is a discount factor, the utility function $\mathcal{Q}(s(\tau))$ is selected with a quadratic form as $\mathcal{Q}(s(t)) = s^T(t)Qs(t)$,

$Q \in \mathbb{R}^3$ is positive definite, and the utility function $U(\mu(\tau, z_i))$ is designed by an integrand positive definite function as [1]

$$\begin{aligned}U(\mu(t, z_i)) &= \int_0^{\mu_1(t, z_i)} [2\lambda_1 \tanh^{-1}(\lambda_1^{-1}v_1)] R_1 dv_1 \\ &\quad + \int_0^{\mu_2(t, z_i)} [2\lambda_2 \tanh^{-1}(\lambda_2^{-1}v_2)] R_2 dv_2 \\ &= \int_0^{\mu(t, z_i)} [2\lambda \tanh^{-T}(\lambda^{-1}v)] R dv\end{aligned}\quad (15)$$

where $\mu(t, z_i) = [\mu_1(t, z_i), \mu_2(t, z_i)]^T$ is the control vector in a RWDA vehicle, $\lambda = \text{diag}(\lambda_1, \lambda_2) > 0$ is the saturating bound matrix of the control vector, $\tanh^{-T}(\cdot) = [\tanh^{-1}(\cdot)]^T$, $v = [v_1, v_2]^T$, and $R = \text{diag}(R_1, R_2) > 0$ is a positive definite matrix.

Definition 1. For the system (11), a state feedback control policy $\mu(t, z_i)$ is called an admissible event-triggered resilient control, if the control $\mu(t, z_i)$ is updated based on an event-triggered mechanism, and can make the infinite horizon integral performance index $J(s(0))$ finite with any initial state $s(0)$, when stabilizing the system dynamics to its origins.

By using the admissible event-triggered resilient control, the cost function in $[t, +\infty)$ can be obtained as

$$\begin{aligned}V(s(t)) &= \int_t^{z_{i+1}} [\gamma^{-\varrho(\tau-t)} \|g(\tau)\|^2 + \mathcal{Q}(s(\tau)) + U(\mu(\tau, z_i))] d\tau \\ &\quad + \sum_{j=i+1}^{+\infty} \int_{z_j}^{z_{j+1}} [\gamma^{-\varrho(\tau-z_j)} \|g(\tau)\|^2 + \mathcal{Q}(s(\tau)) + U(\mu(\tau, z_j))] d\tau,\end{aligned}\quad (16)$$

where time $t \in [z_i, z_{i+1})$.

To compute the optimal control, one can obtain the Hamiltonian function as

$$\begin{aligned}H(V, \mu, s) &= \nabla V^T(s(t)) (\bar{f}(t) + g(t)\mu(t, z_i) + \sigma(t)) \\ &\quad - \varrho\|\delta_g\| + \|g(t)\|^2 + \mathcal{Q}(s(t)) + U(\mu(t, z_i))\end{aligned}\quad (17)$$

where $\|\delta_g\| = \int_t^\infty \gamma^{-\varrho(\tau-t)} \|g(\tau)\|^2 d\tau$ is a constant, and $\nabla V(s(t))$ denotes the partial derivative of $V(s(t))$ with respect to the state $s(t)$.

Under the event-triggered sampling mechanism, the optimal cost function at every triggering instant $t = z_i$, $i \in \mathcal{N}$ yields:

$$\begin{aligned}V^*(s(z_i)) &= \\ &\min_{\mu(t, z_i)} \left[\int_{z_i}^{z_{i+1}} [\gamma^{-\varrho(\tau-z_i)} \|g(\tau)\|^2 + \mathcal{Q}(s(\tau)) + U(\mu(\tau, z_i))] d\tau + \right. \\ &\quad \left. \min_{\mu(\cdot)} \sum_{j=i+1}^{+\infty} \int_{z_j}^{z_{j+1}} [\gamma^{-\varrho(\tau-z_j)} \|g(\tau)\|^2 + \mathcal{Q}(s(\tau)) + U(\mu(\tau, z_j))] d\tau \right]\end{aligned}\quad (18)$$

which satisfies the following HJB equation:

$$\begin{aligned}\nabla V^{*T}(s(z_i)) &\left(\bar{f}(t) + g(t)\mu^*(t, z_i) + \sigma(t) \right) \\ &\quad - \varrho\|\delta_g\| + \|g(t)\|^2 + \mathcal{Q}(s(t)) + U(\mu^*(t, z_i)) = 0.\end{aligned}\quad (19)$$

Note that in the general time-driven optimal control, the HJB equation needs to hold for all time instants, and the control policy $\mu(t)$ is tuned along with the time evolution, as well as the system state $s(t)$ synchronous. In the event-triggered control design, the control policy is only updated at

some triggering instants, but the system states and dynamics are updated in real time. Then, the HJB equation is only required to hold at every triggering instant under a specific event-triggered sampling mechanism.

Lemma 1. *By using the integrand utility function $U(\mu(\tau, z_i))$ in the cost function, the event-triggered resilient optimal control for any time $t \in [z_i, z_{i+1})$ can be derived by the form:*

$$\mu^*(t, z_i) = -\lambda \tanh\left(\frac{1}{2}(\lambda R)^{-1}g^T(z_i)\nabla V^*(s(z_i))\right) \quad (20)$$

where $\mu^*(t, z_i) = [\mu_1^*(t, z_i), \mu_2^*(t, z_i)]^T$, and the control policy is bounded in the sense that $\|\mu_1^*(t, z_i)\| \leq \lambda_1$ and $\|\mu_2^*(t, z_i)\| \leq \lambda_2$.

Proof. Based on the Bellman's optimality principle, we take the derivative of HJB equation (19) with respect to control $\mu(t, z_i)$ at time $t = z_i$, when the control policy is updated, there is

$$\begin{aligned} & \frac{\partial \left(\nabla V^{*T}(s(z_i))(\bar{f}(t) + g(t)\mu(t, z_i) + \sigma(t)) \right)}{\partial \mu(t, z_i)} \\ &= \nabla V^{*T}(s(z_i))g(t) + 2\lambda \tanh^{-T}(\lambda^{-1}\mu(t, z_i))R \\ &= 0. \end{aligned} \quad (21)$$

which indicates

$$2\lambda \tanh^{-1}(\lambda^{-1}\mu(t, z_i)) = -R^{-1}g^T(t)\nabla V^*(s(z_i)).$$

Thus, the event-triggered resilient optimal control can be computed by:

$$\begin{aligned} \mu^*(t, z_i) &= \arg \min_{\mu(t, z_i)} \left(\int_{z_i}^{z_{i+1}} [\gamma^{-\varrho(\tau-z_i)} \|g(\tau)\|^2 \right. \\ &\quad \left. + \mathcal{Q}(s(\tau)) + U(\mu(\tau, z_i))\right] d\tau + V^*(s(z_{i+1})) \Big) \quad (22) \\ &= -\lambda \tanh\left(\frac{1}{2}(\lambda R)^{-1}g^T(z_i)\nabla V^*(s(z_i))\right). \end{aligned}$$

for any time $t \in [z_i, z_{i+1})$, which is constrained as $\|\mu_1^*(t, z_i)\| \leq \lambda_1$ and $\|\mu_2^*(t, z_i)\| \leq \lambda_2$. Thus the proof is completed. $\square$

Then, the optimal cost function at any time t can be further presented as

$$\begin{aligned} V^*(s(t)) &= \int_t^{z_{i+1}} [\gamma^{-\varrho(\tau-t)} \|g(\tau)\|^2 + \mathcal{Q}(s(\tau)) + U(\mu^*(\tau, z_i))\right] d\tau \\ &+ \min_{\mu(\cdot)} \sum_{j=i+1}^{+\infty} \int_{z_j}^{z_{j+1}} [\gamma^{-\varrho(\tau-z_j)} \|g(\tau)\|^2 + \mathcal{Q}(s(\tau)) + U(\mu(\tau, z_j))\right] d\tau. \end{aligned} \quad (23)$$

B. The Event-Triggered Optimal Control Design and Stability Analysis

Based on the event-triggered sampling mechanism, we will develop the event-triggered optimal control policy $\mu^*(t, z_i)$ in driving the RWDA vehicle to track the desired reference trajectory. The architecture of event-triggered resilient optimal control is proposed in the Fig. 3, and the event-triggered condition will be designed as follows.

Fig. 3. The Architecture of Event-Triggered Resilient Optimal Control Mechanism

Lemma 2. *Let $\mu^*(s(t))$ be the optimal state-feedback control policy at time t with the form*

$$\mu^*(s(t)) = -\lambda \tanh\left(\frac{1}{2}(\lambda R)^{-1}g^T(s(t))\nabla V^*(s(t))\right), \quad (24)$$

then, one has $\|\mu^(s(t)) - \mu^*(s(z_i))\| \leq \bar{\lambda} \|s(t) - s(z_i)\|$, where z_i is an arbitrary triggering instant, $\bar{\lambda} > 0$ is a constant.*

Proof. Consider the hyperbolic tangent function $\tanh(\cdot)$, which is continuous, monotone and bounded, and its derivative is $\tanh'(\cdot) = 1 - \tanh^2(\cdot)$. Thus, according to the hyperbolic tangent function's properties, there is

$$\frac{\mu^*(s(t)) - \mu^*(s(z_i))}{s(t) - s(z_i)} \leq \bar{\lambda}, \quad (25)$$

for some positive constant $\bar{\lambda}$, which indicates that $\|\mu^*(s(t)) - \mu^*(s(z_i))\| \leq \bar{\lambda} \|s(t) - s(z_i)\|$.

This gives the proof. $\square$

Theorem 2. *Consider the autonomous driving system and the event-triggered sampling mechanism, the tracking dynamic (11) is asymptotically stable with the event-triggered resilient optimal control $\mu^*(t, z_i)$, when the event-triggered condition for any time $t \in [z_i, z_{i+1})$, $i \in \mathcal{N}$, is given by*

$$\begin{aligned} \mathcal{F}(s(t)) &= \frac{\rho_{\min}(Q)}{\bar{\lambda} \|L_1(t)\|} \|s(t)\|^2 + E(t) \\ &\leq \|s(t) - s(z_i)\| = s_T(t, z_i), \end{aligned} \quad (26)$$

where $E(t) = \frac{\frac{1}{2}\|g(t)\|^2 + U(\mu^*(t))}{\bar{\lambda} \|L_1(t)\|}$, $L_1(t) = \nabla V^{*T}(s(t))g(t)$, and $\rho_{\min}(\cdot)$ means the minimal eigenvalue of a matrix. Besides, the Zeno behavior is strictly ruled out for the vehicle's control process.

Proof. Firstly, we consider the tracking error dynamic of the vehicle with the event-triggering mechanism (26).

Let's select the cost function $V^*(s(t))$ with the event-triggered resilient optimal control $\mu^*(t, z_i)$ as a Lyapunov function, and derive it with respect to time, thus, there is

$$\dot{V}^*(s(t)) = \nabla V^{*T}(s(t))(\bar{f}(t) + g(t)\mu^*(t, z_i) + \sigma(t)). \quad (27)$$

As we pointed out above, the HJB equation can hold for every time with the time-driven optimal control $\mu^*(t)$, thus, it becomes

$$\begin{aligned} \nabla V^{*T}(s(t)) & \left(\bar{f}(t) + g(t)\mu^*(t) + \sigma(t) \right) \\ & - \varrho \|\delta_g\| + \|g(t)\|^2 + \mathcal{Q}(s(t)) + U(\mu^*(t)) = 0. \end{aligned} \quad (28)$$

Note that the discount factor $\varrho > 0$ and can be selected as $\varrho = \frac{\min(\|g(t)\|^2)}{2\|\delta_g\|}$, and based on the vehicle system (11), there is $\min(\|g(t)\|^2) \geq 1$.

Inserting the HJB equation (28) into (27), and based on the Lemma 2, we have

$$\begin{aligned} \dot{V}^*(s(t)) & = \nabla V^{*T}(s(t)) \left(\bar{f}(t) + g(t)\mu^*(t) + \sigma(t) \right) \\ & \quad - L_1(t)\mu^*(t) + L_1(t)\mu^*(t, z_i) \\ & \leq \varrho \|\delta_g\| - \|g(t)\|^2 - \mathcal{Q}(s(t)) - U(\mu^*(t)) \\ & \quad - L_1(t)[\mu^*(t) - \mu^*(t, z_i)] \\ & \leq -\frac{1}{2}\|g(t)\|^2 - U(\mu^*(t)) - \rho_{\min}(Q)\|s(t)\|^2 \\ & \quad + \|L_1(t)\|\bar{\lambda}\|s(t) - s(z_i)\|. \end{aligned} \quad (29)$$

Then it can be found that $\dot{V}^*(s(t)) \leq 0$ for all time $t \in [z_i, z_{i+1}]$, $i \in \mathcal{N}$, when the event-triggering condition (26) holds.

Secondly, we prove that the Zeno behavior is strictly ruled out for the vehicle's control process.

The proof is given by contradiction. Suppose Zeno behavior occurs for the vehicle. Further, without loss of generality, suppose the tracking system (11) exhibits Zeno behavior. Then, there is a finite time $T > 0$, such that $z_i < T$ and $\lim_{l \rightarrow \infty} z_l = T$. To make the contradiction, we will first find that for any $i \in \mathcal{N}$, there is a positive integer $\varepsilon_i > 0$ such that $z_i + \varepsilon_i \leq z_{i+1}$. Consider the the function (29), when the event is triggered as

$$\mathcal{F}(s(t)) \leq s_T(t, z_i) \quad (30)$$

at the time instant is $t = z_i$, then, it yields $\|s(t) - s(z_i)\| = 0$ (the term $s_T(t, z_i)$ is forced to zero), and the error state $s(t)$ will change and drop as

$$\dot{V}^*(s(t)) = -\frac{1}{2}\|g(t)\|^2 - \mathcal{Q}(s(t)) - U(\mu^*(t)) < 0. \quad (31)$$

After that, along with the change of error $s(t)$, when the time $t \in [z_i, z_{i+1}]$, there is

$$\mathcal{F}(s(t)) > s_T(t, z_i).$$

It means that the term $\|L_1(t)\|\bar{\lambda}\|s(t) - s(z_i)\|$ in the time interval $t \in [z_i, z_{i+1}]$ will increase from 0 to $\frac{1}{2}\|g(t)\|^2 + U(\mu^*(t)) + \rho_{\min}(Q)\|s(t)\|^2 > 0$ to trigger the condition (26) again. As we known, the term $\|g(t)\| \geq 1 > 0$ in the vehicle for all time, thus, there must exist a time interval $\varepsilon_i > 0$, such that $z_i + \varepsilon_i \leq z_{i+1}$. Thus, $z_i < T$, $T = \lim_{l \rightarrow \infty} z_l = z_i + \sum_{l=i}^{\infty} \varepsilon_l = +\infty$, which contradicts that T is a finite time. The proof is completed. $\square$

Remark 1. In this paper, the event-triggered mechanism (26) is proposed to effectively balance the frequency/changes in

adjusting the vehicle's control during the running process: **(a)** At the vehicle's urgent tracking stage, the error state $s(t)$ is very large, then the condition (26) is triggered with a shorter time interval, and the event-triggered control can make the vehicle tracking the desired reference faster. **(b)** At the vehicle's steady tracking stage, the error state $s(t)$ can be in some small ranges, then the condition (26) can be triggered with a longer time interval, where the designed control can be un-changed, and no operation need to be changed in the vehicle system.

Compared with the conventional time-based driving approaches, the event-triggered mechanism can reduce much energy and resources in the vehicle's physical operation. And the new event-triggered control can achieve a good balance between the tracking and convenient for the autonomous vehicle.

IV. A SINGLE CRITIC NETWORK BASED EVENT-TRIGGERED CONTROL ALGORITHM

In this section, we design the adaptive resilient event-triggered control algorithm for RWDA vehicles by combining the event-triggered sampling mechanism and the adaptive dynamic programming method, where a single critic network structure is proposed.

Firstly, to obtain the control parameters, we design the single critic network to approximate the optimal cost function based on the higher-order approximation theorem [42] by

$$V^*(s(t)) = \mathcal{W}^T \Phi(s(t)) + \varepsilon(t) \quad (32)$$

where $\mathcal{W} = [\omega_1, \omega_2, \dots, \omega_n]^T \in \mathbb{R}^{n \times 1}$ is the weight vector, n is the number of activation functions, $\Phi(s(t)) = [\phi_1(s(t)), \phi_2(s(t)), \dots, \phi_n(s(t))]^T \in \mathbb{R}^{n \times 1}$ is the independent activation function vector, $\varepsilon(t)$ is the approximated error, which satisfies $\|\varepsilon(t)\| \leq \varepsilon_c$, ε_c can be an arbitrarily small positive constant, when the number of activation functions n is large enough. It can be guaranteed that $\varepsilon(t) \rightarrow 0$ as $n \rightarrow \infty$, based on the approximation theorem [42]. Besides, in the single network framework, we also use the network for tuning the control policy.

Let $V^{(k)}(s(t)) = \mathcal{W}^{(k)T} \Phi(s(t))$ be the k -th approximated value of $V^*(s(t))$. Then, we design the novel adaptive resilient event-triggered control, as presented in Algorithm 1, for the autonomous driving system of vehicles.

Note that in Algorithm 1, the learning procedure is developed based on the policy iteration learning process, and the well-known least square method is applied to update the weight parameters for policy evaluation as

$$\begin{aligned} & \underbrace{\mathcal{W}^{(k)T}}_{\mathcal{X}} \cdot \underbrace{\Phi(s(t)) \left(\bar{f}(t) + g(t)\mu^{(k)}(t) + \sigma(t) \right)}_{\mathcal{Y}} \\ & = \underbrace{\varrho \|\delta_g\| - \|g(t)\|^2 - \mathcal{Q}(s(t)) - U(\mu^{(k)}(t))}_{\mathcal{Z}} \end{aligned} \quad (35)$$

where $\mathcal{W}^{(k)T}$ can be solved by $\mathcal{X} = \mathcal{Z}\mathcal{Y}^T(\mathcal{Y}\mathcal{Y}^T)^{-1}$. Besides, it has $\mathcal{X} = \mathcal{Z}\mathcal{Y}^{-1}$ if the generated data make $\mathcal{Y}$ a full-rank square matrix. Then, we insert the event-triggered sampling

Algorithm 1 Adaptive Resilient Event-Triggered Control Algorithm for the Autonomous Driving System of Vehicles

Initialization

Select an initial admissible control $\mu^{(1)}(t)$, set $k = 1$, $\epsilon > 0$ is a small number and propagate the tracking procedure along with time $t \in [0, +\infty)$

Learning Procedure

- 1: **while** $\|V^{(k+1)} - V^{(k)}\| > \epsilon$ **do**
- 2: (Policy Evaluation) Solve $V^{(k)}(s) = \mathcal{W}^{(k)T} \Phi(s)$ from the Bellman's equation:

$$(\nabla V^{(k)})^T(s(t)) \left(\bar{f}(t) + g(t)\mu^{(k)}(t) + \sigma(t) \right) - \varrho \|\delta_g\| + \|g(t)\|^2 + \mathcal{Q}(s(t)) + U(\mu^{(k)}(t)) = 0 \quad (33)$$

- 3: (Policy Improvement) Tune the control $\mu^{(k)}(t)$ by using:

$$\mu^{(k+1)}(t) = -\lambda \tanh \left(\frac{1}{2} (\lambda R)^{-1} g^T(t) \nabla V^{(k)}(s(t)) \right) \quad (34)$$

- 4: Set $k = k + 1$
- 5: **end while**

Controlling Procedure

- 1: **for** $i \in \mathcal{N}$ **do**
- 2: Run the autonomous driving system (11) and use the event-triggered resilient optimal control policy $\mu^*(t, z_i)$ for the time $t \in [z_i, z_{i+1})$, which has the same parameters with the feedback control $\mu^{(k)}(t)$ obtained from the iteration learning processes (33) and (34)
- 3: **if** the system state $s(t)$ satisfies the designed triggering condition (26):

$$\mathcal{F}(s(t)) \leq s_T(t, z_i)$$

then

- 4: Update the control policy for driving the autonomous vehicle as $\mu^*(t, z_{i+1})$ for the time $t \in [z_{i+1}, z_{i+2})$
- 5: Set $i = i + 1$
- 6: **end if**
- 7: **end for**

End Procedures

mechanism into the controlling procedure of the autonomous driving system, and achieve a new control scheme from Algorithm 1.

Let's suppose the number of activation functions n is selected large enough, and we will provide the convergence analysis of the learning procedure shown in Algorithm 1 in the following theorem.

Theorem 3. Suppose that the control policy $\mu^{(k)}(t)$ is updated according to the equation (34), the cost function $V^{(k)}(s(t))$ is obtained by solving equation (33) with the boundary condition $V^{(k)}(\mathbf{0}) = 0$, where $\mathbf{0}$ is the zero vector, $k = 0, 1, \dots$, and the optimal cost function $V^*(s(t)) \geq 0$ is smooth on a compact domain of validity $\Omega(s)$. Then, the sequence $\{V^{(k)}(s(t)), k = 0, 1, \dots\}$ is monotonically non-increasing, i.e., $V^*(s(t)) \leq$

$V^{(k+1)}(s(t)) \leq V^{(k)}(s(t))$, for all $s(t) \in \Omega(s)$. Moreover, we have

$$\begin{cases} V^{(k)}(s(t)) \rightarrow V^*(s(t)) \\ \mu^{(k)}(s(t)) \rightarrow \mu^*(s(t)) \end{cases} \quad (36)$$

when $k \rightarrow +\infty$ for any state $s(t) \in \Omega(s)$.

Proof. Firstly, let us consider the difference between cost functions $V^{(k+1)}(s(t))$ and $V^{(k)}(s(t))$ along with the system dynamic under control $\mu^{(k+1)}(s(t))$, $\forall s(t)$, we have

$$\begin{aligned} & V^{(k+1)}(s(t)) - V^{(k)}(s(t)) \\ &= \int_t^{+\infty} (\nabla V^{(k+1)})^T(s(\tau)) \left(\bar{f}(\tau) + g(\tau)\mu^{(k+1)}(\tau) + \sigma(\tau) \right) d\tau \\ &\quad - \int_t^{+\infty} (\nabla V^{(k)})^T(s(\tau)) \left(\bar{f}(\tau) + g(\tau)\mu^{(k+1)}(\tau) + \sigma(\tau) \right) d\tau. \end{aligned} \quad (37)$$

According to the policy evaluation equation (33), there are

$$\begin{aligned} & (V^{(k+1)})^T(s(t)) \left(\bar{f}(t) + \sigma(t) \right) \\ &= -(V^{(k+1)})^T(s(t)) g(t) \mu^{(k+1)}(t) + \varrho \|\delta_g\| - \|g(t)\|^2 \\ &\quad - \mathcal{Q}(s(t)) - U(\mu^{(k+1)}(t)) \end{aligned}$$

and

$$\begin{aligned} & (V^{(k)})^T(s(t)) \left(\bar{f}(t) + \sigma(t) \right) \\ &= -(V^{(k)})^T(s(t)) g(t) \mu^{(k)}(t) + \varrho \|\delta_g\| - \|g(t)\|^2 \\ &\quad - \mathcal{Q}(s(t)) - U(\mu^{(k)}(t)), \end{aligned}$$

which make equation (37) to be

$$\begin{aligned} & V^{(k+1)}(s(t)) - V^{(k)}(s(t)) \\ &= \int_t^{+\infty} (\nabla V^{(k+1)})^T(s(\tau)) g(\tau) \mu^{(k+1)}(\tau) d\tau \\ &\quad - \int_t^{+\infty} [(\nabla V^{(k+1)})^T(s(\tau)) g(\tau) \mu^{(k+1)}(\tau) - \varrho \|\delta_g\| \\ &\quad + \|g(\tau)\|^2 + \mathcal{Q}(s(\tau)) + U(\mu^{(k+1)}(\tau))] d\tau \\ &\quad - \int_t^{+\infty} (\nabla V^{(k)})^T(s(\tau)) g(\tau) \mu^{(k+1)}(\tau) d\tau \\ &\quad + \int_t^{+\infty} [(\nabla V^{(k)})^T(s(\tau)) g(\tau) \mu^{(k)}(\tau) - \varrho \|\delta_g\| \\ &\quad + \|g(\tau)\|^2 + \mathcal{Q}(s(\tau)) + U(\mu^{(k)}(\tau))] d\tau. \end{aligned} \quad (38)$$

Based on the definition (15) and using equation (34), it will become:

$$\begin{aligned} & V^{(k+1)}(s(t)) - V^{(k)}(s(t)) \\ &= - \int_t^{+\infty} \left[-\varrho \|\delta_g\| + \|g(\tau)\|^2 + \mathcal{Q}(s(\tau)) + U(\mu^{(k+1)}(\tau)) \right] d\tau \\ &\quad + \int_t^{+\infty} \left[-(\nabla V^{(k)})^T(s(\tau)) g(\tau) \mu^{(k+1)}(\tau) \right. \\ &\quad \left. + (\nabla V^{(k)})^T(s(\tau)) g(\tau) \mu^{(k)}(\tau) \right. \\ &\quad \left. - \varrho \|\delta_g\| + \|g(\tau)\|^2 + \mathcal{Q}(s(\tau)) + U(\mu^{(k)}(\tau)) \right] d\tau \end{aligned} \quad (39)$$

which can be further derived as

$$\begin{aligned}
& V^{(k+1)}(s(t)) - V^{(k)}(s(t)) \\
&= - \int_t^{+\infty} [U(\mu^{(k+1)}(\tau)) - U(\mu^{(k)}(\tau))] d\tau \\
&\quad + \int_t^{+\infty} [2\lambda \tanh^{-T}(\lambda^{-1}\mu^{(k)}(\tau))R(\mu^{(k+1)}(\tau) - \mu^{(k)}(\tau))] d\tau \\
&= - \int_t^{+\infty} \left[\int_{\mu^{(k)}}^{\mu^{(k+1)}} [2\lambda \tanh^{-T}(\lambda^{-1}v)] R dv \right. \\
&\quad \left. - 2\lambda \tanh^{-T}(\lambda^{-1}\mu^{(k)}(\tau))R[\mu^{(k+1)}(\tau) - \mu^{(k)}(\tau)] \right] d\tau
\end{aligned} \tag{40}$$

According to the first integration mid-value theorem, we have

$$\begin{aligned}
& \int_{\mu^{(k)}}^{\mu^{(k+1)}} [2\lambda \tanh^{-T}(\lambda^{-1}v)] R dv \\
&= 2\lambda \tanh^{-T}(\lambda^{-1}\bar{\mu})R[\mu^{(k+1)} - \mu^{(k)}]
\end{aligned} \tag{41}$$

where the elements in $\bar{\mu}$ are between the corresponding elements in $\mu^{(k)}$ and $\mu^{(k+1)}$. As the function $\tanh^{-1}(\cdot)$ is a monotone increasing function, it yields

$$\begin{cases} [\varphi(\lambda^{-1}\bar{\mu})]_i \geq [\varphi(\lambda^{-1}\mu^{(k)})]_i, & [\mu^{(k+1)}]_i \geq [\mu^{(k)}]_i \\ [\varphi(\lambda^{-1}\bar{\mu})]_i \leq [\varphi(\lambda^{-1}\mu^{(k)})]_i, & [\mu^{(k+1)}]_i \leq [\mu^{(k)}]_i \end{cases} \tag{42}$$

where function $\varphi(\cdot) = \tanh^{-T}(\cdot)$ and $[\cdot]_i$ means the i th element of a vector. Thus, there is

$$V^{(k+1)}(s(t)) - V^{(k)}(s(t)) \leq 0, \tag{43}$$

and based on the definition of optimal cost function, we have

$$V^*(s(t)) \leq V^{(k+1)}(s(t)) \leq V^{(k)}(s(t)) \tag{44}$$

for any state $s(t) \in \Omega(s)$, which indicates that the sequence $\{V^{(k)}(s(t)), k = 0, 1, \dots\}$ is monotonic non-increasing.

Secondly, since $\Omega(s)$ is a compact set, according to the Dini's theorem [43], the monotonic non-increasing sequence $\{V^{(k)}(s(t)), k = 0, 1, \dots\}$ will uniform pointwise converge to the optimal cost function $V^*(s(t))$ as $V^{(k)}(s(t)) \rightarrow V^*(s(t))$, along with $k \rightarrow +\infty$.

Finally, from the optimal cost function, the uniform convergence of the control sequence $\{\mu^{(k)}(t), k = 0, 1, \dots\}$ is also achieved during the iteration learning procedure as $\mu^{(k)}(s(t)) \rightarrow \mu^*(s(t))$, along with $k \rightarrow +\infty$.

The proof is thus completed. $\square$

V. SIMULATION CASE STUDIES

We now apply our approach to a simulated autonomous driving system (10) of vehicles, where the modified RWDA vehicle's dynamic function [39] becomes

$$\dot{s}(t) = f(s(t)) + g(s(t))(\mu(t) + u_r(t)) + \sigma(t) \tag{45}$$

where $s(t) = [x_e(t), y_e(t), \theta_e(t)]^T$, $\mu(t) = \begin{bmatrix} w_x(t) \\ v_x(t) \end{bmatrix}$,

$$f(t) = \begin{bmatrix} v_r(t) \cos(\theta_e(t)) \\ v_r(t) \sin(\theta_e(t)) \\ w_r(t) \end{bmatrix}, \quad g(t) = \begin{bmatrix} y_e(t) & -1 \\ -x_e(t) - d_r & 0 \\ -1 & 0 \end{bmatrix},$$

$\sigma(t) = \alpha(t) + g(t)\beta(t)$, $d_r = 1.2$ (m) is the direct distance from the mass center to the rear axle in the vehicle, $\lambda = \text{diag}(\lambda_1, \lambda_2) > 0$ is the saturating bound matrix of the control vector, $\lambda_1 = \lambda_2 = 2$, and we adopt the uncertain signals as shown in [41] with sensor attacks $\alpha(t) = -(0.75 + 0.15 \sin(2.5t))$, $t \geq 0$, and actuator attacks $\beta(t) = [1, 1]^T 0.005 \cos(2.5t) + [0.1 \cos(2t), 0.5 \sin(t)]^T 0.2 \sin(x_e(t)) \cos(y_e(t))$, $t \geq 0$.

Fig. 4. The Evolution of Parameters in Weight Vector $\mathcal{W}$

The initial system state of a desired reference is selected as $[x_r(0), y_r(0), \theta_r(0)]^T = [0, 0, 0]^T$, the desired longitudinal velocity and yaw angular velocity are selected as $v_r(t) = 0.5$ (m/s) and $w_r(t) = 0$ (rad/s) in the reference policy $u_r(t)$. The parameters in the cost function (23) are chosen as $\gamma = 2$, $\varrho = 0.5||g(t)||^2/||\delta_g||$, positive definite matrices in utility functions are $Q = \text{diag}(20, 20, 20)$ and $R = \text{diag}(1, 1)$. And the activation function vector of optimal cost function is selected as $\Phi(s) = [s_1^4, s_2^4, s_3^4, s_1^2 s_2^2, s_1^2 s_3^2, s_2^2 s_3^2, s_1^2 s_2 s_3, s_1 s_2^2 s_3, s_1 s_2 s_3^2, s_1^3 s_2, s_1^3 s_3, s_1 s_3^3, s_2^3 s_3, s_2 s_3^3, \sin(s_1), \sin(s_2), \sin(s_3), \cos(s_1), \cos(s_2), \cos(s_3)]^T$, and the initial weight vector is $\mathcal{W}^{(0)} = \text{rands}(21, 1)$. Then, the evolution of the weight vector is presented in the Fig. 4, where the parameters are get converged during the learning procedure as shown in Algorithm 1.

Fig. 5. The X-Y Diagram of Driving Trajectories from Two Autonomous Vehicles

To verify the availability of the designed resilient event-triggered control scheme, we simulate two autonomous driving systems of RWDA vehicles at different initial system states selected as $[-1.2, 1.2, 0.5]^T$ for the first vehicle, and $[1.2, -1.2, -0.5]^T$ for the second vehicle. The X-Y diagram of driving trajectories is shown in the Fig. 5, where we can find that the two autonomous vehicles are driven to approach the reference under the proposed event-triggered control method, although the uncertain attacking signals occur on actuator and sensor.

Besides, Fig. 5 also shows the advantage of designed event-triggered control, as we have pointed out in Remark 1. Then the corresponding event-triggering evolution processes of the sampling mechanism are depicted in following Fig. 6, where the triggering functions $\mathcal{F}(s(t))$ and $s_T(t, z_i)$ are given by the condition (26). The sampling period of the designed algorithm is also shown in the figure, which displays every sampling instant, and when the event (26) is triggered, the triggering function $s_T(t, z_i)$ will be forced to zero as desired.

Fig. 6. The Evolution of the Triggering Condition and Sampling Period

Along with the procedures in Algorithm 1, the developed event-triggering control policy is obtained, which is determined according to the event-triggering condition (26) and the achieved parameters of weight vector. Fig. 7 displays the event-driven control input that the autonomous vehicle receives in the driving process, which contains the event-driven resilient tracking control policy $\mu(t, z_i)$, and the reference policy $u_r(t)$. As the reference policy $u_r(t) = [0, 0.5]^T$, the obtained vehicle's control inputs are constrained and satisfy the saturating bound condition $u_1(t, z_i) \in [-2, 2]$ and $u_2(t, z_i) \in [-1.5, 2.5]$ for all time instants. Moreover, the comparison between event-triggered and time-driven control trajectories are also presented in Fig. 7 clearly, in which the changing/adjusting times of the vehicle's control with the designed event-triggered method is very smaller than the time-driven method during the vehicle's running process.

Additionally, the tracking error trajectories of the first vehi-

Fig. 7. The Event- and Time-Driven Control Inputs for Vehicles

Fig. 8. The Tracking Error Trajectory under Uncertain Attacking Signals of the Autonomous Driving System

cle are presented in the Fig. 8, where the tracking errors from the autonomous driving system can be found approaching to a region around zero under the proposed event-triggered resilient tracking control as predicted by theorems, and although the uncertain attacking signals occur on actuator and sensor, it further demonstrates the effectiveness of the designed method.

VI. CONCLUSIONS

In this paper, we have designed an adaptive resilient event-triggered control method for an autonomous driving system of RWDA vehicles. Based on the kinematic equation, the tracking control objective was proposed and DoS attacks are considered in the driving dynamic system. An event-triggering condition was presented, which leads to a specific event-triggering mechanism that determines when the control policy was updated, and can effectively balance the frequency/changes in adjusting the vehicle's control during the running process. To overcome the challenge of DoS attacks injected in the tracking system dynamics, the new adaptive resilient control algorithm was

designed based on an iterative single critic learning framework, which can also reduce the control operations of autonomous vehicles. Specific parameter updating strategies were introduced to determine the control policy and event-triggering instants. Finally, the simulation result has clearly shown the effectiveness of our proposed method.

REFERENCES

- [1] M. Abu-Khalaf, F.L. Lewis, "Nearly optimal control laws for nonlinear systems with saturating actuators using a neural network HJB approach," *Automatica*, vol. 41, no. 5, pp. 779-791, May 2005.
- [2] D. Liu, D. Wang, H. Li, "Decentralized stabilization for a class of continuous-time nonlinear interconnected systems using online learning optimal control approach," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 25, no. 2, pp. 418-428, Feb. 2014.
- [3] H. He, S. Chen, K. Li, X. Xu, "Incremental learning from stream data," *IEEE Trans. Neural Netw.*, vol. 22, no. 12, pp. 1901-1914, Dec. 2011.
- [4] H. Zhang, K. Zhang, Y. Cai, J. Han, "Adaptive fuzzy fault-tolerant tracking control for partially unknown systems with actuator faults via integral reinforcement learning method," *IEEE Trans. Fuzzy Syst.*, vol. 27, no. 10, pp. 1986-1998, Oct. 2019.
- [5] C. Mu, Z. Ni, C. Sun, H. He, "Air-breathing hypersonic vehicle tracking control based on adaptive dynamic programming," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 28, no. 3, pp. 584-598, Mar. 2017.
- [6] J. Biggs, W. Holderbaum, "Optimal kinematic control of an autonomous underwater vehicle," *IEEE Trans. Autom. Contr.*, vol. 54, no. 7, pp. 1623-1626, Jul. 2009.
- [7] B. Kiumarsi, F.L. Lewis, Z.P. Jiang, " H_∞ control of linear discrete-time systems: Off-policy reinforcement learning," *Automatica*, vol. 78, pp. 144-152, Apr. 2017.
- [8] T. Bian, Y. Jiang, Z.P. Jiang, "Adaptive dynamic programming for stochastic systems with state and control dependent noise," *IEEE Trans. Autom. Contr.*, vol. 61, no. 12, pp. 4170-4175, Dec. 2016.
- [9] Q. Wei, D. Liu, Q. Lin, R. Song, "Discrete-time optimal control via local policy iteration adaptive dynamic programming," *IEEE Trans. Cybern.*, vol. 47, no. 10, pp. 3367-3379, Oct. 2017.
- [10] F.L. Lewis, K.G. Vamvoudakis, "Reinforcement learning for partially observable dynamic processes: adaptive dynamic programming using measured output data," *IEEE Trans. Syst. Man Cybern. B Cybern.*, vol. 41, no. 1, pp. 14-25, Feb. 2011.
- [11] H. He, X. Zhong, "Learning without external reward," *IEEE Comp. Int. Mag.*, vol. 13, no. 3, pp. 48-54, Aug. 2018.
- [12] Z. Ni, H. He, X. Zhong, D.V. Prokhorov, "Model-free dual heuristic dynamic programming," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 26, no. 8, pp. 1834-1839, Aug. 2015.
- [13] K. Zhang, H. Zhang, Y. Cai, R. Su, "Parallel optimal tracking control schemes for mode-dependent control of coupled Markov jump systems via integral RL method," *IEEE Trans. Autom. Sci. Engin.*, to be published, doi:10.1109/TASE.2019.2948431, 2019.
- [14] E.B. Kosmatopoulos, A. Kouvelas, "Large scale nonlinear control system fine-tuning through learning," *IEEE Trans. Neural Netw.*, vol. 20, no. 6, pp. 1009-1023, Jun. 2012.
- [15] H. Zhang, L. Cui, X. Zhang, Y. Luo, "Data-driven robust approximate optimal tracking control for unknown general nonlinear systems using adaptive dynamic programming method," *IEEE Trans. Neural Netw.*, vol. 22, no. 12, pp. 2226-2236, Dec. 2011.
- [16] Y. Park, M. Choi, K.Y. Lee, "An optimal tracking neuro-controller for nonlinear dynamic systems," *IEEE Trans. Neural Netw.*, vol. 7, no. 5, pp. 1099-1110, Sep. 1996.
- [17] H. Modares, F.L. Lewis, "Optimal tracking control of nonlinear partially-unknown constrained-input systems using integral reinforcement learning," *Automatica*, vol. 50, no. 7, pp. 1780-1792, Jul. 2014.
- [18] K. Zhang, H. Zhang, Y. Mu, C. Liu, "Decentralized tracking optimization control for partially unknown fuzzy interconnected systems via reinforcement learning method," *IEEE Trans. Fuzzy Syst.*, to be published, doi:10.1109/TFUZZ.2020.2966418, 2020.
- [19] B. Luo, D. Liu, T. Huang, D. Wang, "Model-free optimal tracking control via critic-only Q-learning," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 27, no. 10, pp. 2134-2144, Oct. 2016.
- [20] D. Wang, C. Mu, D. Liu, 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, Apr. 2018.
- [21] X. Yang, H. He, "Adaptive critic designs for event-triggered robust control of nonlinear systems with unknown dynamics," *IEEE Trans. Cybern.*, vol. 49, no. 6, pp. 2255-2267, Jun. 2019.
- [22] L. Dong, X. Zhong, C. Sun, H. He, "Event-triggered adaptive dynamic programming for continuous-time systems with control constraints," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 28, no. 8, pp. 1941-1952, Aug. 2017.
- [23] D. Wang, M. Ha, J. Qiao, "Self-learning optimal regulation for discrete-time nonlinear systems under event-driven formulation," *IEEE Trans. Autom. Contr.*, vol. 65, no. 3, pp. 1272-1279, Mar. 2020.
- [24] C. Zhao, J. He, Q.G. Wang, "Resilient distributed optimization algorithm against adversarial attacks," *IEEE Trans. Autom. Contr.*, to be published, doi:10.1109/TAC.2019.2954363, 2020.
- [25] H. Modares, B. Kiumarsi, F.L. Lewis, F. Ferrese, A. Davoudi, "Resilient and robust synchronization of multiagent systems under attacks on sensors and actuators," *IEEE Trans. Cybern.*, vol. 50, no. 3, pp. 1240-1250, Mar. 2020.
- [26] Z. Wang, Y. Wu, L. Liu, H. Zhang, "Adaptive fault-tolerant consensus protocols for multiagent systems with directed graphs," *IEEE Trans. Cybern.*, vol. 50, no. 1, pp. 25-35, Jan. 2020.
- [27] C. Chen, F.L. Lewis, S. Xie, H. Modares, Z. Liu, S. Zuo, A. Davoudi, "Resilient adaptive and H_∞ controls of multi-agent systems under sensor and actuator faults," *IEEE Trans. Autom. Contr.*, vol. 102, pp. 19-26, Apr. 2019.
- [28] D. Meng, K.L. Moore, "Studies on resilient control through multiagent consensus networks subject to disturbances," *IEEE Trans. Cybern.*, vol. 44, no. 11, pp. 2050-2064, Nov. 2014.
- [29] Y. Tang, D. Zhang, P. Shi, W. Zhang, F. Qian, "Event-based formation control for multi-agent systems under DoS attacks," *IEEE Trans. Autom. Contr.*, to be published, doi:10.1109/TAC.2020.2979936, 2020.
- [30] Y. Yuan, Z. Wang, P. Zhang, H. Liu, "Near-optimal resilient control strategy design for state-saturated networked systems under stochastic communication protocol," *IEEE Trans. Cyber.*, vol. 48, no. 8, pp. 3155-3167, Aug. 2019.
- [31] E. Mousavinejad, X. Ge, Q.L. Han, F. Yang, L. Vlacic, "Resilient tracking control of networked control systems under cyber attacks," *IEEE Trans. Cyber.*, to be published, doi:10.1109/TCYB.2019.2948427, 2019.
- [32] M. Xue, Y. Tang, W. Ren, F. Qian, "Practical output synchronization for asynchronously switched multi-agent systems with adaption to fast-switching perturbations," *Automatica*, vol. 116, no. 118917, Jun. 2020.
- [33] J. Hu, J. Shen, D. Lee, "Resilient stabilization of switched linear control systems against adversarial switching," *IEEE Trans. Autom. Contr.*, vol. 62, no. 8, pp. 3820-3834, Aug. 2017.
- [34] J. Dong, Y. Wu, G.H. Yang, "Reliable state feedback control of T-S fuzzy systems with sensor faults," *IEEE Trans. Fuzzy Syst.*, vol. 23, no. 2, pp. 421-433, Apr. 2015.
- [35] M. Zhu, S. Martinez, "On the performance analysis of resilient networked control systems under replay attacks," *IEEE Trans. Autom. Contr.*, vol. 59, no. 3, pp. 804-808, Mar. 2014.
- [36] J. Dong, Y. Wu, G.H. Yang, "A new sensor fault isolation method for T-S fuzzy systems," *IEEE Trans. Cyber.*, vol. 47, no. 9, pp. 2437-2447, Sep. 2017.
- [37] Y. Yuan, P. Zhang, Z. Wang, L. Guo, "On resilient strategy design of multi-tasking optimal control for state-saturated systems with nonlinear disturbances: The time-varying case," *Automatica*, vol. 107, pp. 138-145, Sep. 2019.
- [38] A.Y. Lu, G.H. Yang, "Observer-based control for cyber-physical systems under denial-of-service with a decentralized event-triggered scheme," *IEEE Trans. Cyber.*, to be published, doi:10.1109/TCYB.2019.2944956, 2019.
- [39] G. Ma, M. Ghasemi, X. Song, "Integrated powertrain energy management and vehicle coordination for multiple connected hybrid electric vehicles," *IEEE Trans. Veh. Technol.*, vol. 67, no. 4, pp. 2893-2899, Apr. 2018.
- [40] Z.P. Jiang, H. Nijmeijer, "Tracking control of mobile robots: A case study in backstepping," *Automatica*, vol. 33, no. 7, pp. 1393-1399, Jul. 1997.
- [41] X. Jin, W.M. Haddad, T. Yucelen, "An adaptive control architecture for mitigating sensor and actuator attacks in cyber-physical systems," *IEEE Trans. Autom. Contr.*, vol. 62, no. 11, pp. 6058-6064, Nov. 2017.
- [42] B.A. Finlayson, *The Method of Weighted Residuals and Variational Principles*, New York, Academic Press, 1990.
- [43] R.G. Bartle, D.R. Sherbert, *Introduction to Real Analysis*, 3rd ed. New York, USA: Wiley, 2000.