Reinforcement Learning-Based Event-Triggered Robust Optimal Control for Mobile Euler-Lagrange Systems with Dead-Zone and Saturation Actuators

Tan-Luy Nguyen[1*], Huu-Toan Tran[1], Trong-Toan Tran[2], and Cong-Thanh Pham[3]
[1] Industrial University of Ho Chi Minh City, Vietnam
[2] Institute of Applied Mechanics and Informatics, Vietnam Academy of Science and Technology, Vietnam
[3] Ho Chi Minh City University of Food Industry, Vietnam
Email: nguyentanluy@iuh.edu.vn, tranhuutoan@iuh.edu.vn, toan1003@gmail.com, phamcongthanh1978@gmail.com

***Abstract*—This paper proposes a reinforcement learning (RL)-based event-triggered robust optimal control method for mobile Euler-Lagrange systems with both dead-zone and saturation from actuators. Firstly, kinematics and dynamics of the system are integrated into the equivalent system, where both of the dead-zone and saturation inputs are treated. Secondly, event-triggered robust optimal control and dead-zone disturbance laws are designed, where their parameters are only updated when a triggering condition occurs. Via RL techniques, the new triggering condition is introduced. The method not only guarantees the stability of the closed system and the convergence of the cost function to the bounded $\mathcal{L}_2$ -gain optimal value but also relaxes identification procedures for unknown nonlinear functions. Additionally, it maintains the minimum inter-event time between two sequent triggering instants greater than zero, thus the Zeno's behavior is avoided. Finally, the simulation of a nonholonomic wheeled mobile robot system with dead-zone and saturated torques is implemented to verify the effectiveness of the proposed method.**

***Index Terms*—euler-lagrange systems, event-triggered control, reinforcement learning, dead-zone and saturation, optimal control**

I. Introduction

In in the recent years, the control design of Euler-Lagrange systems has received much attention due to the practical application ability, such as mobile robots [1], autonomous vehicles [2]. For the design, the system dynamics model is formulated by Euler-Lagrange equations [1]-[4], where identifying the correct model of nonholonomic constraint for the mobile systems is a major challenge. The second challenge is that the design not only takes into account the dynamics but also kinematics [1], [2]. The dynamics always contains disturbances such as external, unmodeled and unstructured uncertain disturbances. Therefore, the main problem of controlling the Euler-Lagrange systems is to design the controllers that provide optimality and reject such disturbances.

Practically, the control signals from the actuators of Euler-Lagrange dynamics are often constrained by the dead-zone phenomenon due to the physical limitation. In [5], the unknown dead-zone is considered and its uncertainty is compensated by a neural network (NN). Inspired by the work in [5], various control methods dealing with the dead-zone phenomenon are studied and reported [6]-[8]. In [9], the slope of dead-zone is defuzzified to a deterministic value and the adaptive control scheme for a robot manipulator is designed via fuzzy logic. The work in [10] separates the dead-zone of a robot manipulator into two parts, which are modeled by bounded disturbance. Unfortunately, the existing control schemes are not considered distributed optimal control with dead-zone inputs. Most recently, our previous work [11] proposes an H_∞ optimal control algorithm for physically interconnected mobile Euler-Lagrange systems with slipping, skidding and dead-zone. The algorithm can reject the dead-zone disturbance and approximate the optimal control, concurrently.

Besides the dead-zone factors, the industrial actuators are also constrained by saturation due to the physical limitations of voltages, currents, flows, and torques, etc. That constraint can make the closed-loop unstable. Literature [12]-[16] consider the input constraints along with external disturbances, where however optimal control methods are not presented. Meanwhile, the work in [17] introduces an algorithm for both optimality and saturation.

The above mentioned studies only consider the dead-zone and the saturation independently of each other. In practice, there exist two kinds of factors in the particular actuators that must be dealt with. To the best of our knowledge, control of systems with both dead-zone and the saturation are rarely considered. Furthermore, the robust optimization control method for such systems is

Manuscript received February 17, 2020; revised February 10, 2021.
This research is funded by the Vietnam National Foundation for Science and Technology Development (NAFOSTED) under grant number 107.04-2019.25.

doi: 10.18178/ijmerr.10.3.116-127

not still introduced. That is the motivation of the work in this paper.

Over the past decade, reinforcement learning (RL) [18] has been emerging as powerful techniques to support control designs. In Euler-Lagrange systems, RL is used to develop optimal controllers [19]. Additionally, RL exploits the zero-sum differential graphical game theory to design real controllers for Euler-Lagrange systems in the manner of distributed robust optimality [1]. RL, namely adaptive dynamic programming (ADP), is also employed to design H_∞ optimal control schemes for systems with external disturbances and saturated-inputs [20]-[22]. In the schemes, there are two policies of two players in a differential game are built, one for the optimal control policy and the other for the worst disturbance policy. These policies are formed if solutions od Hamilton-Jacobi-Issac (HJI) equation is solved. Unfortunately, the solution of the HJI equation can not be found by analytic.

Most recently, event-triggering mechanisms [23] have been employed for control design [24], [25]. Compared with the traditional event-time control methods based on periodical sample time, event-triggered controllers have many advantages since they only update parameters and send the control signal to the plants when a new event occurs. Hence, the burden of computation bandwidth and communication is significantly overcome (see an overview in [26]). In [24], event-triggered tracking control of Euler-Lagrange kinematics was researched. Narayanan and Jagannathan [25] introduce event-triggered distributed optimal control for affine nonlinear interconnected systems while Zhu et al. [27] propose event-triggered optimal control for saturated-input systems.

By the above analytics, the event-triggering mechanism has not been considered the phenomenon of both dead-zone and saturation in actuators of the Euler-Lagrange systems for optimal control. To obtain a comprehensive solution, we first propose a novel RL-based event-triggered robust optimal control method for Euler-Lagrange systems with dead-zone and saturation actuators. The main contribution of this paper includes three aspects. 1) Kinematics and dynamics of the system are integrated into the equivalent system, where both dead-zone and saturation inputs are treated. 2) Event-triggered robust optimal control and dead-zone disturbance laws are designed, where their parameters are only updated when a triggering condition occurs. The event-triggering condition is designed such that the Zeno's behavior is excluded. 3) The proposed method avoids the identification procedure and guarantees the stability of the closed system along with the convergence of the value function to the bounded $\mathcal{L}_2$-gain optimal value.

The rest of this paper is organized as follows: Section II describes the preliminary, system dynamics and event-triggered feedforward control. Section III provides a design of RL-based event-triggered robust optimal control with input constraint of dead-zone and saturation. The simulation is conducted in Section IV, and Section V gives a brief conclusion.

II. Preliminary, System Dynamics and Event-Triggered Feed-Forward Control

A. Notation and Definition

Notation: $\mathbb{R}$, $\mathbb{R}^n$, $\mathbb{R}^{n\times m}$ are the set of real numbers, the n-dimensional Euclidean space, and the set of all real $n\times m$ matrices, respectively. The symbol $\|.\|$ denotes the vector or matrix norm in $\mathbb{R}^n$ or $\mathbb{R}^{n\times m}$. The superscript $\top$ is used for the transpose. I_n denotes a n-dimensional identify matrix. $\mathrm{Diag}(x)$ is a diagonal matrix. $\mathcal{L}_2[0,\infty)$ is the Banach space if there exists $\forall d(t)\in\mathcal{L}_2[0,\infty)$ then $\int_0^\infty \|d(t)\|^2 d\tau \in \mathcal{L}_2[0,\infty)$. $\lambda_{\min}(.)$, $\lambda_{\max}(.)$ denote the minimum and maximum eigenvalues, respectively. Note that the dimensions of all matrices in the paper are assumed to be compatible if they are not shown explicitly.

Definition 1 (UUB [28]): The equilibrium point x_0 of system $\dot{x}=f(x,u), x\in\mathbb{R}^n$ is said to be uniformly ultimately bounded (UUB) if there exists a compact set $\Omega\in\mathbb{R}^n$ so that for all $x_0\in\Omega$, there exists a bound B and a time $T(B,x_0)$ such that $\|x-x_0\|\le B$ for all $t>t_0+T$.

B. System Dynamics

Considering a model of nonholonomic mobile Euler-Lagrange systems, of which the kinematics is written by

$$\dot{q}=J(q)v \tag{1}$$

where $q=[q_1,q_2,\ldots,q_n]^\top\in\mathbb{R}^n$ is a vector of generalized coordinates, $v\in\mathbb{R}^{n-m}$ is a velocity vector. The full-rank matrix $J(q)\in\mathbb{R}^{n\times(n-m)}$ induces from a set of smooth and linearly independent vector fields spanning the null space of an associated-constraint matrix $A(q)\in\mathbb{R}^{m\times n}$ that holds $J^\top(q)A^\top(q)=0$ [1].

Using the Euler-Lagrange formulation, system dynamics is presented as

$$M(q)\ddot{q}+C(q,\dot{q})\dot{q}+G(q)+\tau_d=BD(\tau)-A^\top\lambda \tag{2}$$

where the inertia matrix $M(q)\in\mathbb{R}^{n\times n}$ is the symmetric positive definite. According to [1], the Coriolis and centrifugal matrix $C(q,\dot{q})\in\mathbb{R}^n$, the gravitational force $G(q)\in\mathbb{R}^n$ and the external disturbance $\tau_d\in\mathbb{R}^n$ are bounded. $B\in\mathbb{R}^{n\times(n-m)}$ is the input transformation matrix, $\lambda\in\mathbb{R}^m$ is the constraint force vector. The function $D(\tau(t)):\mathbb{R}^{n-m}\mapsto\mathbb{R}^{n-m}$ (Fig. 1), where $\tau(t)\in\mathbb{R}^{n-m}$ is the control torques, denotes the dead-zone factor, i.e.,

$$\vartheta(t)=D(\tau)=\begin{cases} m^r(\tau-b^r), & b^r\le\tau\le\bar{\tau}\\ 0, & -b^l<\tau<b^r\\ m^l(\tau+b^l), & -\bar{\tau}\le\tau\le -b^l\\ D_{sat}, & \tau>\bar{\tau}\\ -D_{sat}, & \tau<-\bar{\tau}\end{cases} \tag{3}$$

where the right and left slopes m^r, m^l of the dead-zone are unknown positive, the breakpoints of the input nonlinearity b^r, b^l are unknown positive, $\bar{\tau}$ is an upper bounded value of the torques due to the saturation actuator, D_{sat} is the saturated dead-zone value. This form is completely different from that found in [4] for robot control with dead-zone inputs.

Equation (3) can be represented as follows:

$$\vartheta(t)=\begin{cases} m\tau-\tau_b & -\bar{\tau}\le\tau\le\bar{\tau}\\ D_{sat}, & \tau>\bar{\tau}\\ -D_{sat}, & \tau<-\bar{\tau}\end{cases} \tag{4}$$

where

$$m=\begin{cases} m^r, 0\le\tau\le\bar{\tau}\\ m^l, -\bar{\tau}\le\tau<0\end{cases},\ \tau_b=\begin{cases} mb^r & b^r\le\tau\le\bar{\tau}\\ m\tau & b^l\le\tau\le b^r\\ -mb^l & -\bar{\tau}\le\tau\le b^l\end{cases}$$

Assumption 1 [4]**:** $\|m\|\le b_m^\star$ with $b_m^\star=\max\{m^r,m^l\}$, $\|\tau_b\|\le\rho^\star$, where $b_m^\star$ and $\rho^\star$ are unknown positive constants.

Note that one can robustly estimates some lower bounds $b_m>0$, $\rho>0$, such that $b_m\le b_m^\star$, $\rho\le\rho^\star$ [29].

Substituting the differentiation in both sides of (1) and (4) to dynamics (2), then multiplying both sides of the yielded result by $J^\top(q)$ one can obtains that

$$\begin{aligned}\dot{v}=&-\bar{M}^{-1}J^\top\left(\bar{M}\dot{J}+C(q,\dot{q})J\right)v-\bar{M}^{-1}J^\top G\ (q)\\ &-\bar{M}^{-1}J^\top\tau_d+\bar{M}^{-1}J^\top B\vartheta\end{aligned} \tag{5}$$

where $\bar{M}(q)=J^\top(q)M(q)J(q)$. To facilitate the control design, a following nonlinear system is derived from (1) and (5):

$$\begin{cases}\dot{x}_q=g_q(x_q)x_v\\ \dot{x}_v=f_v(x_q,x_v)+g_v(x_q,x_v)\vartheta+k_v(x_q,x_v)d_v\\ y=x_q\end{cases} \tag{6}$$

where y is the system output, $x_q=q, x_v=v$, $d_v=\tau_d$, $\bar{C}=\bar{M}\dot{J}+C(q,\dot{q})J$, $k_v(q,v)=-\bar{M}^{-1}J^\top$, $g_q(x_q)=J(q)$, $g_v(x_q,x_v)=\bar{M}^{-1}J^\top B$, $f_v(x_q,x_v)=-\bar{M}^{-1}J^\top(\bar{C}v+G)$.

Remark 1: It is shown in [1] that for practical robotic systems, $g_q(x_q)$, $g_v(x_q,x_v)$, $k_v(x_q,x_v)$ are nominally known since the nominal matrices J,B and M are well defined.

To facilitate the later control design, the following lemma, assumptions and definition are introduced.

Lemma 1 [1]: There exist positive constants b_M, b_J and b_C such that $\|M^{-1}\|\le b_M$, $\|J\|\le b_J$, $\|C\|\le b_C$.

Assumption 2: From Lemma 1, one has $\|g_q\|\le b_{gq}$, $\|g_v\|\le b_{gv}$, $\|k_v\|\le b_{kv}$, where b_{gq},b_{gv}, b_{kv} are known positive constants. Furthermore, for unknown positive scalars b_{fv}, b_{dv}, one has $\|f_v(x_q,x_v)\|\le b_{fv}$, $\|d_v\|\le b_{dv}$.

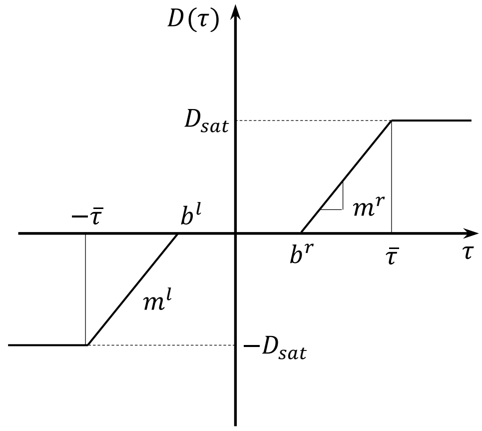

Figure 1. Dead-zone and saturation in actuators.

C. Event-Triggered Feed-Forward Control

In [30], the separate kinematics and dynamics can be integrated into an equivalent system to design a robust optimal control scheme. However, the scheme burdens computational bandwidth. In the paper, we therefore propose another scheme using an event-triggering mechanism to remove the drawback.

Firstly, the coordinate of (6) is changed as

$$\begin{cases} z_q(t)=y(t)-y_0(t)\\ z_v(t)=x_v(t)-x_{vd}(t)\end{cases} \tag{7}$$

where $y_0(t)$ is the reference trajectory, x_{vd} is the virtual control signal such that $x_{vd}=x_{vd}^a+x_{vd}^\star$, where x_{vd}^a is the feedforward control input and $x_{vd}^\star$ is the feedback control input.

Assumption 3: The reference trajectory $y_0(t)$ is smooth and bounded.

Now, the event-triggering mechanism is introduced. Define a monotonically increasing sequence of aperiodic instants $\{t_0,t_1,...,t_k,t_{k+1},...\}$, where $t_k, k=0,1,...$ is a triggering instant on a certain event. If the states are sampled by $\underline{x}_h=x_h(t_k)(h=q,v)$ at t_k, the triggering errors between the current states and the sampled states are

$$e_h = x_h - \underline{x}_h, t_k \le t < t_{k+1} \tag{8}$$

It is easy to know $e_h(t_k) = 0$ at $t_k, k = 0,1,\ldots$. Similarly, sampling $\underline{z}_h = z_h\left(t_k\right)$ and using (7), (8) we obtain

$$e_h = z_h - \underline{z}_h, t_k \le t < t_{k+1} \tag{9}$$

Then, the event-triggered virtual control law $\underline{x}_{vd} = x_{vd}(t_k)$ and the event-triggered actual control law $\underline{\vartheta} = \vartheta(t_k)$ at a triggering instant $t_k, k = 0,1,\ldots$, are given by

$$\begin{cases} \underline{x}_{vd} = \underline{x}_{vd}^a + \underline{x}_{vd}^\star \\ \underline{\vartheta} = \underline{\vartheta}^a + \underline{\vartheta}^\star \end{cases} \tag{10}$$

where $\underline{x}_{vd}^\star$, $\underline{\vartheta}^\star$ are the event-triggered feedback control laws optimizing a cost function will be designed later. $\underline{x}_{vd}^a = x_{vd}^a(t_k)$, $\underline{\vartheta}^a = \vartheta^a(t_k)$ are the event-triggered feedforward control laws updated at t_k using $\underline{z}_h$ and kept unchanged through a zero-order holder (ZOH) until the next event occurs at t_{k+1}. $\underline{x}_{vd}^a$ and $\underline{\vartheta}^a$ are designed as follows:

$$\begin{cases} g_q(x_q)\underline{x}_{vd}^a = \dot{\underline{y}}_0 - \chi \underline{z}_q \\ g_v(x_q, x_v)\underline{\vartheta}^a = \dot{\underline{x}}_{vd} - \chi \underline{z}_v \end{cases} \tag{11}$$

where $\dot{\underline{y}}_0 = \dot{y}_0(t_k)$, $\dot{\underline{x}}_{vd} = \dot{x}_{vd}(t_k)$, and χ is the positive gain. The following Lemma is introduced to show the effectiveness of the event-triggered feedforward control.

Lemma 2: Let the system dynamics be (6), the event-triggered control laws be (10), where the the event-triggered feedforward control laws $\underline{x}_{vd}^a$ and $\underline{\vartheta}^a$ are designed in (11), and the robust optimal control laws $\underline{x}_{vd}^\star$ and $\underline{\vartheta}^\star$ are assumed to stabilize the closed system:

$$\begin{bmatrix} \dot{z}_q \\ \dot{z}_v \end{bmatrix} = \begin{bmatrix} 0 \\ f_v \end{bmatrix} + \begin{bmatrix} g_q & 0 \\ 0 & g_v \end{bmatrix}\begin{bmatrix} \underline{x}_{vd}^\star \\ \underline{\vartheta}^\star \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & k_v \end{bmatrix}\begin{bmatrix} 0 \\ d_v \end{bmatrix} \tag{12}$$

Suppose that the triggering condition satisfies the following inequality

$$\left\|e_h\right\| \le \kappa \left\|z_h\right\|, \forall h = q, v \tag{13}$$

With $0 < \kappa \le b_g / \chi$, $\chi \ge (2b_g + 1/4)$, $b_g = \max(b_{gq}, b_{gv})$. Then, the control problem of system (6) is transformed to the event-triggered distributed control problem of system (12).

Proof: Choose the Lyapunov function as

$$\mathcal{J} = \frac{1}{2}\left(z_q^\top z_q + z_v^\top z_v\right) \tag{14}$$

Taking the time derivative of $\mathcal{J}$ along with (6) using (7)- (11), one obtains

$$\begin{aligned} \dot{\mathcal{J}} = & -\sum_{h=q,v} \chi z_h^\top \underline{z}_h + \sum_{h=q,v} z_h^\top g_h z_{h+1} \\ & + \begin{bmatrix} z_q^\top \\ z_v^\top \end{bmatrix}^\top \left(\begin{bmatrix} 0 \\ f_v \end{bmatrix} + \begin{bmatrix} g_q & 0 \\ 0 & g_v \end{bmatrix}\begin{bmatrix} \underline{x}_{vd}^\star \\ \underline{\vartheta}^\star \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & k_v \end{bmatrix}\begin{bmatrix} 0 \\ d_v \end{bmatrix} \right) \end{aligned} \tag{15}$$

where $z_{q+1} = z_v$, $z_{v+1} = 0$. The first term is computed based on (9):

$$-\sum_{h=q,v} \chi z_h^\top \underline{z}_h = -\sum_{h=q,v} \chi z_h^\top z_h + \sum_{h=q,v} \chi z_h^\top e_{i,h} \tag{16}$$

The second term in (15) is transformed as

$$\begin{aligned} \sum_{h=q,v} z_h^\top g_h z_{h+1} & \le \frac{1}{2}\sum_{h=q,v} z_h^\top \left\|g_h\right\| z_h + \frac{1}{2}\sum_{h=q,v} z_{h+1}^\top \left\|g_h\right\| z_{h+1} \\ & \le \sum_{h=q,v} z_h^\top \left\|g_h\right\| z_h \end{aligned} \tag{17}$$

Substituting (16) , (17) into (15) we have

$$\begin{aligned} \dot{\mathcal{J}} \le & \begin{bmatrix} z_q^\top \\ z_v^\top \end{bmatrix}^\top \left(\begin{bmatrix} 0 \\ f_v \end{bmatrix} + \begin{bmatrix} g_q & 0 \\ 0 & g_v \end{bmatrix}\begin{bmatrix} \underline{x}_{vd}^\star \\ \underline{\vartheta}^\star \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & k_v \end{bmatrix}\begin{bmatrix} 0 \\ d_v \end{bmatrix} \right) \\ & -(\chi - b_g - \frac{1}{4})\sum_{h=q,v} z_h^\top z_h + \sum_{h=q,v} \chi z_h^\top e_h \end{aligned} \tag{18}$$

Using the event-triggering condition (13) yields

$$\begin{aligned} \dot{\mathcal{J}} \le & \begin{bmatrix} z_q^\top \\ z_v^\top \end{bmatrix}^\top \left(\begin{bmatrix} 0 \\ f_v \end{bmatrix} + \begin{bmatrix} g_q & 0 \\ 0 & g_v \end{bmatrix}\begin{bmatrix} \underline{x}_{vd}^\star \\ \underline{\vartheta}^\star \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & k_v \end{bmatrix}\begin{bmatrix} 0 \\ d_v \end{bmatrix} \right) \\ & - \sum_{h=q,v} \left(\chi(1-\kappa) - b_g - \frac{1}{4} \right)\left\|z_h\right\|^2 \\ \le & \begin{bmatrix} z_q^\top \\ z_v^\top \end{bmatrix}^\top \left(\begin{bmatrix} 0 \\ f_v \end{bmatrix} + \begin{bmatrix} g_q & 0 \\ 0 & g_v \end{bmatrix}\begin{bmatrix} \underline{x}_{vd}^\star \\ \underline{\vartheta}^\star \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & k_v \end{bmatrix}\begin{bmatrix} 0 \\ d_v \end{bmatrix} \right) \end{aligned} \tag{19}$$

It can be seen that if there exists control inputs that stabilize the closed system (12), the first term in the right-hand side of (19) becomes negative. Then, according to Definition 1, the tracking error $z = \left[z_q^\top, z_v^\top\right]^\top$ is UUB.

The Proof is completed.

Note $\underline{\vartheta} = \underline{\vartheta}^a + \underline{\vartheta}^\star$ and recall (4), we decompose the dead-zone disturbance such that $\underline{\tau}_b = \underline{\tau}_b^a + \underline{\tau}_b^\star$. The following lemma shows the constraints of $\underline{\vartheta}^a$ and $\underline{\tau}_b$.

Lemma 3: Given the feedforward input $\underline{\vartheta}^a$ and feedback control input $\underline{\vartheta}^\star$, let $\underline{\vartheta}^a$ and $\underline{\tau}_b^a$ be constrained by

$$\begin{cases} \left\|\underline{\tau}_b^a\right\| \le \rho(1 - \tanh(1)), -\overline{\tau} \le \tau \le \overline{\tau} \\ \left\|\underline{\vartheta}^a\right\| \le D_{sat} - \delta \tanh(1), \text{othewise} \end{cases} \tag{20}$$

and $\underline{\tau}_b^\star$, $\underline{\vartheta}^\star$ be approximated by the hyperbolic tangent function that is used to map $\mathbb{R}$ onto the intervals

$(-\rho,\rho)$ and $(-\delta,\delta)$. Then, there exist $\rho \in \mathbb{R}^+$, $\delta \in \mathbb{R}^+$, $\delta \leq D_{sat}$ such that $\|\underline{\tau}_b\| \leq \rho$ and $\|\underline{\vartheta}\| \leq D_{sat}$.

Proof: Motivated by [31], one can approximate $\underline{\tau}_b^\star(? = \rho\tanh(?/\rho$, $\underline{\vartheta}^\star(\circ) = \delta\tanh(\circ/\delta)$ where $\|\bullet\| \leq \rho$, $\|\circ\| \leq \delta$, and utilize (20) respectively along with the monotonic property of the functions tanh for completing the proof.

If the minimum interval between events from the current to the next $t_{\min} = \min_k(t_{k+1} - t_k)$ reduces to zero, the Zeno's behavior will happen [32]. As a result, the calculation will overload. The following theorem shows that the Zeno's behavior is excluded.

Theorem 1: When using the event-triggered feedforward control laws in Lemma 2, the Zeno's behavior is avoided as the minimum inter-event time is lower bounded by a nonzero positive number.

Proof: The proof follows from [17], [33] and is omitted here.

III. Event-Triggered Robust Optimal Control with Input Constraint Based on Reinforcement Learning

After designing the event-triggered feedforward control laws to transform (6) to (12), this section designs the robust optimal control laws.

To facilitate the design, we rewrite the tracking dynamics (12) as

$$\dot{z} = f_z(t) + g(q,v)\bar{m}u^\star - g(q,v)\psi^\star + k(q,v)d \tag{21}$$

where $f_z(t) = [0, f_v^\top]^\top$, $u^\star = [\underline{x}_{vd}^{\star\top}, \tau^{\star\top}]^\top$, $d = [0, d_v^\top]^\top$, $\psi^\star = [0, \tau_b^{\star\top}]^\top$, $\tau^\star$ is actual optimal control input, $\tau_b^\star$ is the worst dead-zone disturbance. $g(q,v) = \mathrm{diag}[g_q, g_v]$, $\bar{m} = \mathrm{diag}[I_{n\times n}, mI_{(n-m)\times(n-m)}]$, $k(q,v) = \mathrm{diag}[0, k_v]$.

For robust optimal control with input constraint, the general disturbances including dead-zone and external disturbances need to be rejected. Inspired by [11], we use RL to design control with disturbance rejection.

For robust control, the following bounded $\mathcal{L}_2$-gain inequality for some attenuation constants $\gamma > \gamma^\star > 0$ is satisfied:

$$\int_0^\infty \left(z^\top Qz + b_{\bar{m}}U(u^\star)\right)dt \leq \int_0^\infty \left(U(\psi^\star) + \gamma^2 d^\top d\right)dt \tag{22}$$

where Q is a positive definite matrix, $b_{\bar{m}} = \|\bar{m}\|$. As $u^\star$, $\psi^\star$ are constrained, inspired by [34], the costs are defined by nonnegative nonquadratic functions $U(u^\star)$ and $U(\psi^\star)$ as

$$U(u^\star) = 2\delta\int_0^{u^\star} \tanh^{-\top}(s/\delta)S\,ds \tag{23}$$

$$U(\psi^\star) = 2\rho\int_0^{\psi^\star} \tanh^{-\top}(s/\rho)R\,ds \tag{24}$$

where R, S are diagonal positive definite matrices. The cost function is then defined as

$$J = \int_0^\infty \left(z^\top Qz + b_{\bar{m}}U(u^\star) - U(\psi^\star) - \gamma^2 d^\top d\right)dt \tag{25}$$

Because the cost function (25) contains the disturbances, the two-player zero-sum game theory [35] is used to seek the optimal value $V^\star$. That is, the control player finds the policy for $u^\star$ to minimize the value while two disturbance players try to find the policies for $\psi^\star$ and $d^\star$ to maximize it:

$$V^\star = \min_{u^\star}\max_{\psi^\star}\max_{d^\star} J \tag{26}$$

If the condition $\min_{u^\star}\max_{\psi^\star}\max_{d^\star} J = \max_{u^\star}\min_{\psi^\star}\min_{d^\star} J$ holds, the saddle point exists. In this case, a minimum positive solution $V^\star > 0$ is the Nash equilibrium value.

In RL, value functions are predefined by state feedback policies. We therefore derive it from (25) as

$$V(z(t)) = \int_t^\infty \left(z^\top Qz + b_{\bar{m}}U(u) - U(\psi) - \gamma^2 d^\top d\right)dt \tag{27}$$

Using the infinitesimal version of (27) subject to (12), the Hamiltonian is given by

$$\begin{aligned} H &= z^\top Qz + b_{\bar{m}}U(u) - U(\psi) - \gamma^2 d^\top d \\ &+ \left(\nabla V^\star\right)^\top \left(f_z + g(q,v)\bar{m}u - g(q,v)\psi + k(q,v)d\right) = 0 \end{aligned} \tag{28}$$

Define the event-triggering errors as

$$\underline{e} = z(t) - \underline{z}(t) \tag{29}$$

where $\underline{z}(t) = z(t_k)$, $t_k \leq t < t_{k+1}$, $\forall k \in \mathbb{N}$. Then, the event-triggered robust optimal control law $\underline{u}^\star$, the event-triggered disturbance law $\underline{\psi}^\star$ and the time-triggered disturbance law $d^\star$ are given by

$$\underline{u}^\star = -\delta\tanh(\underline{M}^\star), \underline{M}^\star = \frac{1}{2\delta}S^{-1}g^\top(\underline{q},\underline{v})\nabla V^\star(\underline{z}) \tag{30}$$

$$\underline{\psi}^\star = -\rho\tanh(\underline{N}^\star), \underline{N}^\star = \frac{1}{2\rho}R^{-1}g^\top(\underline{q},\underline{v})\nabla V^\star(\underline{z}) \tag{31}$$

$$d^\star = \frac{1}{2\gamma^2}k^\top(q,v)\nabla V^\star(z) \tag{32}$$

where $\nabla V^\star \triangleq \partial V^\star/\partial z$.

Using (30)-(32) for (28), the event-triggering HJI equation is obtained as

$$H = z^\top Qz + b_{\bar{m}}U(u) - U(\psi) - \gamma^2 d^\top d + \left(\nabla V^\star\right)^\top \left(f_z + g(q,v)\bar{m}u - g(q,v)\psi + k(q,v)d \right) = 0 \tag{33}$$

The HJI (33) cannot be solved analytically. Thus, an NN-based event-triggered control scheme is designed to approximate its solutions.

Given a compact set $z(t) \in \Omega$, the smooth solution $V^\star(z(t)) \in C^1(\Omega)$ can be approximated by a NN as [34]

$$V^\star(z) = W^\top \phi(z) + \varepsilon(z) \tag{34}$$

$$\nabla V^\star(z) = \nabla\phi^\top(z) W + \nabla\varepsilon(z) \tag{35}$$

where $\phi(z)$ is the activation function of $\hbar$ neuron units in the hidden layer, $W \in \mathbb{R}^\hbar$ are the ideal weights, ∇ is a partial derivative operator, and $\varepsilon(z)$ is the approximation error. If $\phi(z)$ is a completely independent basis set, then $\|\phi(z)\| \le b_\phi$, $\|\nabla\phi(z)\| = \|\partial\phi(z)/\partial z\| \le b_{\nabla\phi}$, $\|\varepsilon(z)\| \le b_\varepsilon$, $\|\nabla\varepsilon(z)\| = \|\partial\varepsilon(z)/\partial z\| \le b_{\nabla\varepsilon}$, where b_ϕ, $b_{\nabla\phi}$, b_ε, $b_{\nabla\varepsilon}$ are positive upper bounds. Additionally, $\nabla\phi(z)$ is Lipschitz for a positive constant $L_{\nabla\phi}$

$$\|\nabla\phi(z) - \nabla\phi(\underline{z})\| \le L_{\nabla\phi}\|z - \underline{z}\| = L_{\nabla\phi}\|\underline{e}\| \tag{36}$$

Then, the NN-based event-triggered HJI equation is given by

$$\underline{z}^\top Q\underline{z} + W^\top \nabla\phi\left(f_z + g\bar{m}\underline{u} - g\underline{\psi} + kd \right) + b_{\bar{m}}U(\underline{u}) - U(\underline{\psi}) - \gamma^2 d^\top d - e_H = 0 \tag{37}$$

where e_H is the HJI approximation error, and

$$\underline{u} = -\delta \tanh\left(\frac{1}{2\delta} S^{-1} g^\top(\underline{q},\underline{v}) \nabla\phi^\top(\underline{z}) W + \nabla\varepsilon(\underline{z}) \right) \tag{38}$$

$$\underline{\psi} = -\rho \tanh\left(\frac{1}{2\rho} R^{-1} g^\top(\underline{q},\underline{v}) \nabla\phi^\top(\underline{z}) W + \nabla\varepsilon(\underline{z}) \right) \tag{39}$$

$$d = \frac{1}{2\gamma^2} k^\top(q,v)\left(\nabla\phi^\top(z) W + \nabla\varepsilon(z) \right) \tag{40}$$

From the function approximation theory [34], $\exists \mathbb{N}(b_{e_H} > 0) : \sup_{\delta\in\Omega}\|e_H\| \le b_{\varepsilon H}$. Furthermore, if $\hbar \to \infty$, e_H converges uniformly to zero.

To avoid identifying the function $f_z(t)$ when online approximating the solutions to (37), the modified version of RL [36] is used, i.e., the integral of HJI equation (37) is taken as

$$\int_{t-T}^{t} \left(W^\top \nabla\phi(z)\left(f_z + g\bar{m}\underline{u} - g\underline{\psi} + kd \right) + r \right) d\tau = \varepsilon_H \tag{41}$$

where $T > 0$, $\varepsilon_H = \int_{t-T}^{t} e_H d\tau$ is the approximation error. For a positive constant $b_{\varepsilon H}$, one has $\|\varepsilon_H\| \le b_{\varepsilon H}$, and

$$r = \underline{z}^\top Q\underline{z} + b_{\bar{m}}U(\underline{u}) - U(\underline{\psi}) - \gamma^2\|d\|^2 \tag{42}$$

As the ideal weights of NN are unknown, the approximation of value function and its partial derivative are approximated by

$$\hat{V} = \hat{W}^\top \phi, \nabla\hat{V} = \nabla\phi^\top \hat{W} \tag{43}$$

Using the event-triggered control law (38), the event-triggered disturbance laws (39) and (40), the time-triggered disturbance laws are approximated by

$$\underline{\hat{u}} = -\delta \tanh(\underline{\hat{M}}),\ \underline{\hat{M}} = \frac{1}{2\delta} S^{-1} g^\top(\underline{q},\underline{v}) \nabla\phi^\top(\underline{z}) \hat{W} \tag{44}$$

$$\underline{\hat{\psi}} = -\rho \tanh(\underline{\hat{N}}),\ \underline{\hat{N}} = \frac{1}{2\rho} R^{-1} g^\top(\underline{q},\underline{v}) \nabla\phi^\top(\underline{z}) \hat{W} \tag{45}$$

$$\hat{d} = \frac{1}{2\gamma^2} k^\top(q,v) \nabla\phi^\top(z) \hat{W} \tag{46}$$

Using (44)-(46) for the dynamics (12) yields

$$\dot{z} = f_z(t) + g(q,v)\bar{m}\underline{\hat{u}} - g(q,v)\underline{\hat{\psi}} + k(q,v)\hat{d} \tag{47}$$

On the other hand, substituting (43) to (41), we have the residual error $e_r \in \mathbb{R}$:

$$e_r = \int_{t-T}^{t} \left(\hat{W}^\top \nabla\phi(z)(f_z + g\bar{m}\hat{u} - g\underline{\hat{\psi}} + k\hat{d}) + \hat{r} \right) d\tau \tag{48}$$

where

$$\hat{r} = \underline{z}^\top Q\underline{z} + b_{\bar{m}}U(\underline{\hat{u}}) - U(\underline{\hat{\psi}}) - \gamma^2\|\hat{d}\|^2 \tag{49}$$

Dynamics (47) can be changed to (50) (see (50) below) by multiplying both sides of (47) with $\nabla\phi(z)$ and then taking the integral:

$$\begin{aligned} &\int_{t-T}^{t} \nabla\phi(z)\left(f_z + (g\bar{m}\underline{\hat{u}} - g\underline{\hat{\psi}} + k\hat{d}) \right) d\tau \\ &= \int_{t-T}^{t} \frac{\partial\phi(z)}{\partial z} \dot{z} d\tau = \phi(z(t)) - \phi(z(t-T)) \\ &= \Delta\phi(z(t)) \end{aligned} \tag{50}$$

The residual error (48) becomes

$$e_r = \hat{W}^\top \Delta\phi + \int_{t-T}^{t} \hat{r} d\tau \tag{51}$$

Now, we propose a NN weight-tuning law that brings $\hat{W}$ to W such that the squared function of the residual error $E = 1/2(e_r^\top e_r) \to 0$ when $t \to \infty$. By using the a gradient-descent method we have

$$\dot{\hat{W}} = -\alpha \frac{\partial E}{\partial \hat{W}} = -\alpha \frac{\partial E}{\partial e_r} \frac{\partial e_r}{\partial \hat{W}} = -\alpha\Delta\phi\left(\Delta\phi^\top \hat{W} + \int_{t-T}^{t} \hat{r} d\tau \right) \tag{52}$$

where $\alpha > 0$ denotes an update rate. The NN approximation error dynamics ($\tilde{W} = W - \hat{W}$) is written as

$$\dot{\tilde{W}} = -\alpha\Delta\phi\left(\Delta\phi^\top \tilde{W} - \varepsilon_H\right) \quad (53)$$

The triggering law for the parameter update of (44) and (45) is designed as

$$\|\underline{e}\|^2 \geq \underline{E}_T = (1-\beta)\frac{(1-\eta)\lambda_{\min}(Q)\|\underline{z}\|^2 + U(\underline{\hat{u}}) - U(\underline{\hat{\psi}}) - \gamma^2\left\|\hat{d}\right\|^2}{\left(\frac{1}{\eta}-1\right)\lambda_{\min}(Q) + \mu\left\|\hat{W}\right\|^2} \quad (54)$$

where $0<\eta,\beta<1$, μ is defined in Appendix. It is worth emphasizing that updating the parameters of (44) and (45) only occur when the squared norm of the event-triggering error exceeds a threshold E_T defined in (54).

The stability and convergence of the proposed control method needs to be proven.

Proof of the stability and convergence: Consider the Lyapunov function candidate:

$$L = L_1 + L_2 + L_3 \quad (55)$$

where

$$\begin{cases} L_1 = V^*(\underline{z}) \\ L_2 = \int_{t-T}^{t} V^*(z(\tau))d\tau \\ L_3 = \frac{1}{2}\int_{t-T}^{t} \operatorname{trace}\left(\tilde{W}^T\tilde{W}\right)d\tau \end{cases} \quad (56)$$

Case 1: The system is within inter-event intervals: $V^\star(\underline{z})$ is unchanged, thus the derivative of L_1 is zero. Taking the derivative of L_2 along the trajectories of (47) and utilizing $V^{\star\top}(z)f_z$ from the optimal HJI equation, ones has

$$\dot{L}_2 = \int_{t-T}^{t} \dot{\bar{L}}_2 d\tau \quad (57)$$

where

$$\begin{aligned} \dot{\bar{L}}_2 &= -z^\top Qz - b_{\bar{m}}U(u^\star) + U(\psi^\star) + \gamma^2 d^{\star\top}d^\star \\ &\quad -\nabla V^{\star\top}\left(g\bar{m}u^\star - g\psi^\star + kd^\star\right) \\ &\quad +\nabla V^{\star\top}\left(g\bar{m}\underline{\hat{u}} - g\underline{\hat{\psi}} + k\hat{d}\right) \end{aligned} \quad (58)$$

Substituting $u^\star$, $\psi^\star$ from (30) to (23) and (31) into (24) one has

$$\begin{aligned} U(u^\star) &= 2\delta\tanh^{-\top}\left(u^\star/\delta\right)Su^\star + \delta^2\bar{S}\ln\left(\bar{1} - \frac{u^{\star 2}}{\delta^2}\right) \\ &= (\nabla V^\star)^\top g\tanh\left(\frac{1}{2\delta}S^{-1}g^\top\nabla V^\star\right) \\ &\quad + \delta^2\bar{S}\ln\left(\bar{1} - \tanh^2\left(\frac{1}{2\delta}S^{-1}g^\top\nabla V^\star\right)\right) \end{aligned} \quad (59)$$

$$\begin{aligned} U(\psi^\star) &= 2\rho\tanh^{-\top}\left(\psi^\star/\rho\right)R\psi^\star + \rho^2\bar{R}\ln\left(\bar{1} - \frac{\psi^{\star 2}}{\rho^2}\right) \\ &= (\nabla V^\star)^\top g\tanh\left(\frac{1}{2\rho}R^{-1}g^\top\nabla V^\star\right) \\ &\quad + \rho^2\bar{R}\ln\left(\bar{1} - \tanh^2\left(\frac{1}{2\rho}R^{-1}g^\top\nabla V^\star\right)\right) \end{aligned} \quad (60)$$

where $\bar{1} = [1,...,1]^\top$, $\bar{S}$, $\bar{R}$ is a row vector containing the main diagonal elements of S, R. Using (30), (31) and (60), $\dot{\bar{L}}_2$ can be rewritten as

$$\begin{aligned} \dot{\bar{L}}_2 &= -z^\top Qz + \gamma^2 d^{\star\top}d^\star \\ &\quad -\nabla V^{\star\top}\left(g\bar{m}u^\star + kd^\star + g\bar{m}\underline{\hat{u}} - g\underline{\hat{\psi}} + k\hat{d}\right) \\ &\quad +\delta^2\bar{S}\ln\left(\bar{1} - \tanh^2(M^\star)\right) - \rho^2\bar{R}\ln\left(\bar{1} - \tanh^2(N^\star)\right) \end{aligned} \quad (61)$$

The terms in (61) can be rewritten as

$$\begin{aligned} \delta^2\bar{S}\ln\left(\bar{1} - \tanh^2(M^\star)\right) &= \int_{\underline{\hat{u}}}^{u^\star} 2\delta\tanh^{-\top}(s/\delta)Sds \\ &\quad + U(\underline{\hat{u}}) - \delta\nabla V^{\star\top}g\tanh(M^\star) \end{aligned} \quad (62)$$

$$\begin{aligned} \rho^2\bar{R}\ln\left(\bar{1} - \tanh^2(N^\star)\right) &= \int_{\underline{\hat{\psi}}}^{\psi^\star} 2\rho\tanh^{-\top}(s/\rho)Rds \\ &\quad + U(\underline{\hat{\psi}}) - \rho\nabla V^{\star\top}g\tanh(N^\star) \end{aligned} \quad (63)$$

$$\nabla V^{\star\top}g\underline{\hat{u}} = \int_{u^\star}^{\underline{\hat{u}}} 2\delta M^{\star\top}Sds - \delta\nabla V^{\star\top}g\tanh(M^\star) \quad (64)$$

$$\nabla V^{\star\top}g\underline{\hat{\psi}} = \int_{\psi^\star}^{\underline{\hat{\psi}}} 2\rho N^{\star\top}Rds - \rho\nabla V^{\star\top}g\tanh(N^\star) \quad (65)$$

$$k^\top\nabla V^\star = 2\gamma^2 d^\star \quad (66)$$

$$2\gamma^2 d^{\star\top}\hat{d} - \gamma^2 d^{\star\top}d^\star \leq \gamma^2\hat{d}^\top\hat{d} \quad (67)$$

$$\begin{aligned} z^\top Qz &= \underline{z}^\top Q\underline{z} - 2\underline{z}^\top Q\underline{e} + \underline{e}^\top Q_{ii}\underline{e} \\ &\geq (1-\eta)\underline{z}^\top Q\underline{z} - \left(\frac{1}{\eta}-1\right)\underline{e}^\top Q\underline{e} \end{aligned} \quad (68)$$

Substituting (62)-(68) into (61) yields

$$\begin{aligned} \dot{\bar{L}}_2 &\leq -(1-\eta)\lambda_{\min}(Q)\|\underline{z}\|^2 + \left(\frac{1}{\eta}-1\right)\lambda_{\min}(Q)\|\underline{e}\|^2 \\ &\quad -U(u^\star) + U(\underline{\hat{\psi}}) + \xi_1 + \xi_2 + \gamma^2\left\|\hat{d}\right\|^2 \end{aligned} \quad (69)$$

where

$$\xi_1 = \int_{\underline{\hat{u}}}^{u^\star} 2\delta\left(\tanh^{-1}(s/\delta) + M^\star\right)^\top S ds,$$

$$\xi_2 = \int_{\underline{\hat{\psi}}}^{\psi^\star} 2\rho\left(\tanh^{-1}(s/\rho) + N^\star\right)^\top R ds.$$

By changing $s = -\rho\tanh(\nu)$, ξ_1 can be presented as

$$\begin{aligned}\xi_1 &\le \int_{\underline{\hat{M}}}^{M^\star} 2\delta^2\left(\nu - M^\star\right)^\top S d\nu \\ &= \delta^2\left(M^\star - \underline{\hat{M}}\right)^\top S\left(M^\star - \underline{\hat{M}}\right) \\ &\le \delta^2\|S\|\left\|M^\star - \underline{\hat{M}}\right\|^2\end{aligned} \tag{70}$$

Similarly, ξ_2 is changed as

$$\begin{aligned}\xi_2 &\le \int_{\underline{\hat{N}}}^{N^\star} 2\rho^2\left(\nu - N^\star\right)^\top R d\nu \\ &= \rho^2\left(N^\star - \underline{\hat{N}}\right)^\top R\left(N^\star - \underline{\hat{N}}\right) \\ &\le \rho^2\|R\|\left\|N^\star - \underline{\hat{N}}\right\|^2\end{aligned} \tag{71}$$

Using $\nabla V^\star(z)$ from (35) for $M^\star$ and $N^\star$ in (30) and (31), $\underline{\hat{M}}$ and $\underline{\hat{N}}$ from (44) and (45) into (69), note that $\tilde{W} = W - \hat{W}$, one obtains

$$\begin{aligned}\xi_1 &\le \delta^2\|S\|\Big\|\frac{1}{2}S^{-1}g^\top(\underline{q},\underline{v})\nabla\phi^\top(\underline{z})\hat{W} - \frac{1}{2}R^{-1}g^\top(q,v)\nabla\phi^\top(z)\hat{W} \\ &\quad - \frac{1}{2}S^{-1}g^\top(q,v)\nabla\phi^\top(z)\left(\tilde{W} + \nabla\varepsilon(z)\right)\Big\|^2 \\ &\le \frac{1}{2}\delta^2\|S\|\left\|S^{-1}\right\|^2\Big(\left\|g^\top(\underline{q},\underline{v})\nabla\phi^\top(\underline{z}) - g^\top(q,v)\nabla\phi^\top(z)\right\|^2\left\|\hat{W}\right\|^2 \\ &\quad + \left\|g^\top(q,v)\nabla\phi^\top(z)\left(\tilde{W} + \nabla\varepsilon(z)\right)\right\|^2\Big)\end{aligned} \tag{72}$$

$$\begin{aligned}\xi_2 &\le \rho^2\|R\|\Big\|\frac{1}{2}R^{-1}g^\top(\underline{q},\underline{v})\nabla\phi^\top(\underline{z})\hat{W} - \frac{1}{2}R^{-1}g^\top(q,v)\nabla\phi^\top(z)\hat{W} \\ &\quad - \frac{1}{2}R^{-1}g^\top(q,v)\nabla\phi^\top(z)\left(\tilde{W} + \nabla\varepsilon(z)\right)\Big\|^2 \\ &\le \frac{1}{2}\rho^2\|R\|\left\|R^{-1}\right\|^2\Big(\left\|g^\top(\underline{q},\underline{v})\nabla\phi^\top(\underline{z}) - g^\top(q,v)\nabla\phi^\top(z)\right\|^2\left\|\hat{W}\right\|^2 \\ &\quad + \left\|g^\top(q,v)\nabla\phi^\top(z)\left(\tilde{W} + \nabla\varepsilon(z)\right)\right\|^2\Big)\end{aligned} \tag{73}$$

Using the inequality $(ab - cd)^2 \le 2a^2(b-d)^2 + 2d^2(a-c)^2$ and Assumption 2 and (36), one yields

$$\begin{aligned}&\left\|g^\top(\underline{q},\underline{v})\nabla\phi^\top(\underline{z}) - g^\top(q,v)\nabla\phi^\top(z)\right\|^2 \\ &\le 2\left\|g(\underline{q},\underline{v})\right\|^2\left\|\nabla\phi(\underline{z}) - \nabla\phi(z)\right\|^2 + 2\left\|\nabla\phi(z)\right\|^2\left\|g(\underline{q},\underline{v}) - g(q,v)\right\|^2 \\ &\le 2b_g^2 L_{\nabla\phi}^2\left\|\underline{e}\right\|^2 + 4b_{\nabla\phi}^2 b_g^2\end{aligned} \tag{74}$$

From (72)-(74), one can rewrite (69) as

$$\begin{aligned}\dot{L}_2 &\le -(1-\eta)\lambda_{\min}(Q)\left\|\underline{z}\right\|^2 - U(\underline{\hat{u}}) + U(\underline{\hat{\psi}}) \\ &\quad + \gamma^2\left\|\hat{d}\right\|^2 + \left(\left(\frac{1}{\eta} - 1\right)\lambda_{\min}(Q) + \mu_2\left\|\hat{W}\right\|^2\right)\left\|\underline{e}\right\|^2 \\ &\quad + \mu_1\left\|\tilde{W}\right\|^2 + \mu\left\|\tilde{W}\right\|\end{aligned} \tag{75}$$

where

$$\mu_1 = \mu_2^2 b_{\nabla\varepsilon}^2\left(\left\|S^{-1}\right\| + \left\|R^{-1}\right\|\right),\ \mu = b_{\nabla\phi}^2 b_g^2\left(\delta^2\left\|S^{-1}\right\| + \rho^2\left\|R^{-1}\right\|\right),$$

Let's take derivative of L_3 along the trajectories of (53), we have

$$\dot{L}_3 = \int_{t-T}^{t}\left(-\alpha\tilde{W}^\top\Pi\tilde{W} + \alpha\tilde{W}^\top\Delta\phi\varepsilon_H\right)d\tau \tag{76}$$

where $\Pi = \Delta\phi\Delta\phi^\top > 0$. Using the Young inequality we have:

$$\dot{L}_3 \le -(\alpha - 1)\lambda_{\min}(\Pi)\int_{t-T}^{t}\left\|\tilde{W}\right\|^2 d\tau \tag{77}$$

Substituting (77) and (75) into (55) we have:

$$\begin{aligned}\dot{L} &< \int_{t-T}^{t}\Big(-(1-\eta)\lambda_{\min}(Q)\left\|\underline{z}\right\|^2 - U(\underline{\hat{u}}) + U(\underline{\hat{\psi}}) \\ &\quad + \gamma^2\left\|\hat{d}\right\|^2 + \left(\left(\frac{1}{\eta} - 1\right)\lambda_{\min}(Q) + \mu\left\|\hat{W}\right\|^2\right)\left\|\underline{e}\right\|^2 \\ &\quad - \mu_2\left(\left(\left\|\tilde{W}\right\| - \frac{\mu}{2\mu_2}\right)^2 + \lambda_1\right)d\tau\end{aligned} \tag{78}$$

where $\mu_2 = \alpha - \mu_1 - 1$. Select $\alpha > \mu_1 + 1$ then $\mu_2 > 0$, $\lambda_1 = \mu^2/(4\mu_2)$.

Applying the triggering condition (54) for (78) one obtains

$$\begin{aligned}\dot{L} &< -\frac{1}{T}\beta\Big((1-\eta)\lambda_{\min}(Q)\left\|\underline{z}\right\|^2 + \lambda_{\min}(S)\left\|\underline{\hat{v}}\right\|^2 \\ &\quad - U(\underline{\hat{\psi}}) - \gamma^2\left\|\hat{d}\right\|^2\Big) < 0, \forall t\end{aligned} \tag{79}$$

if only if

$$\left\|\tilde{W}\right\| \ge b_{\tilde{W}} = \sqrt{\lambda_1/\mu_2} + \mu/2\mu_2 \tag{80}$$

Thus, the closed dynamics is UUB stable. Note that by properly adjusting α, R, S, it can be made $b_{\tilde{W}}$ be a considerably small value.

Case 2: In the case of the system at event-triggering instants, $\forall t = t_k, \forall k \in \mathbb{N}$. Taking the difference of the Lyapunov function candidate (35) one has

$$\Delta L = V^{\star}(\underline{z}(t_k)) - V^{\star}(\underline{z}(t_{k-1})) + \int_{t_k - T}^{t_k} V^{\star}(\underline{z}(t_k))d\tau$$
$$-\int_{t_k^- - T}^{t_k^-} V^{\star}(z(t^-))d\tau + \frac{1}{2}\tilde{W}^T(t_k)\tilde{W}(t_k) \tag{81}$$
$$-\frac{1}{2}\tilde{W}^T(t^-)\tilde{W}(t^-)$$

From (78) since $\dot{L} < 0$ and the state of the system along with the function approximation are continuous, it can be shown that

$$\int_{t_k - T}^{t_k} V^{\star}(\underline{z}(t_k))d\tau \le \int_{t_k^- - T}^{t_k^-} V^{\star}(z(t^-))d\tau \tag{82}$$

$$\int_{t_k - T}^{t_k} V^{\star}(\underline{z}(t_k))d\tau \le \int_{t_k^- - T}^{t_k^-} V^{\star}(z(t^-))d\tau \tag{83}$$

Then, ΔL becomes

$$\begin{aligned} \Delta L &\le V^{\star}(\underline{z}(t_k)) - V^{\star}(\underline{z}(t_{k-1})) \\ &\le V^{\star}(z(t^-)) - V^{\star}(\underline{z}(t_{k-1})) \\ &\le -K^{\star}\left\| z(t^-) - \underline{z}(t_{k-1}) \right\| = -K^{\star}\left\| \underline{e}(t_{k-1}) \right\| \end{aligned} \tag{84}$$

where $K^{\star}$ belongs to a class-κ function [37]. Thus, Lyapunov function (84) is still decreasing at arbitrary triggering instant $t = t_k, k \in \mathbb{N}$.

From (84) and (79), it can be seen that the closed tracking error dynamics is asymptotically stable.

The Proof is completed.

IV. Simulation Study

Consider a wheeled mobile robot (WMR) defined in [1]. The state vectors and parameters of WMR are $q = [X, Y, \Theta]^{\top}$, $v = [\Upsilon, \Omega]^{\top}$, where X, Y are the point of the robot centered position on the global Cartesian coordinate system, Ω is the direction, Υ and Ω are the linear and rotational velocities, and

$$S(q) = \begin{bmatrix} \cos(\Omega) & 0 \\ \sin(\Omega) & 0 \\ 0 & 1 \end{bmatrix}, M = \begin{bmatrix} m & 0 \\ 0 & I_1 \end{bmatrix}, B = \begin{bmatrix} \frac{1}{r_1} & \frac{b_1}{r_1} \\ \frac{1}{r_1} & -\frac{b_1}{r_1} \end{bmatrix}$$

where m, I_1 denote the value of the mass and the moment of inertia of the platform, motors and wheels, respectively, with $r_1 = 0.05m$ $b_1 = 0.5m$, $m = 10kg$, $I_1 = 5kg.m^2$. One assumes that the robust optimal control torques of the right and left actuators are saturated by $\left\|\tau_l^{\star}\right\| \le \|\delta\|$ and $\left\|\tau_r^{\star}\right\| \le \|\delta\|$, where the indexes $\{l,r\}$ denote the left and right torques, respectively, and $\|\delta\| = 1.5\text{N.m}$. The torques are also affected by the dead-zone actuators with $m^r = m^l = 1.25$, $b_1^r = 0.85$, $b_1^l = 1.0$, $b_2^r = 1.0$, $b_2^l = 0.85$, $b_{\overline{m}} = 2$. The external disturbance is $\tau_d = 0.05\text{rand(t)}$.

The smooth desired eight-shaped trajectory of the virtual robot for tracking $q_d = [X_d, Y_d, \Omega_d]^{\top}$ is generated by the velocities v_{rd}:

$$v_{rd} = \begin{bmatrix} \Upsilon_{rd} \\ \Omega_{rd} \end{bmatrix} = \begin{pmatrix} \sqrt{\cos^2 t + 4\cos^2(2t)} \\ (2\sin t\cos(2t) - 4\sin(2t)\cos t)/(\cos^2 t + 4\cos^2(2t)) \end{pmatrix}$$

It is initial at $q_d(0) = [0.1, -0.6, \pi/6]^{\top}$.

The weight vector of NN is defined as $\hat{W} = [\hat{W}_1, \hat{W}_2, \ldots, \hat{W}_{15}]^{\top}$ whose initial values are zeros. The adaptive gains are selected as $\alpha = 100$. The activation function vector of critic NN with 15 elements is chosen as $\phi(z) = [z_X^2, z_X z_Y, z_X z_\Theta, z_X z_\Upsilon, z_X z_\Omega, z_Y^2, z_Y z_\Theta,$ $z_Y z_\Upsilon, z_Y z_\Omega, z_\Theta^2, z_\Theta z_\Upsilon, z_\Theta z_\Omega, z_\Upsilon^2, z_\Upsilon z_\Omega, z_\Omega^2]^{\top}$ $R = I_{4\times 4}$, $\gamma = 5$, $Q = I_{5\times 5}$, $S = I_{4\times 4}$ $\gamma = 5$. The initial position and velocities of WMR are $q(0) = [0.5, -0.5, 0]^{\top}$, $v(0) = [0, 0]^{\top}$, respectively. The integral interval T is chosen as $0.01s$. The parameters of the triggering condition is $\beta = 0.2$, $\eta = 0.8$, $\mu = 10$.

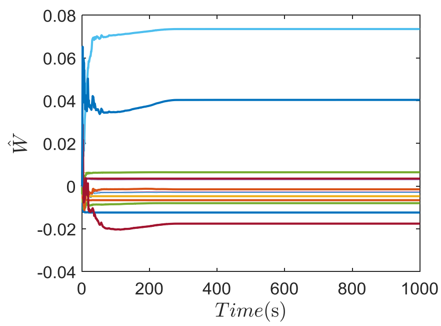

Figure 2. Convergence of NN weights.

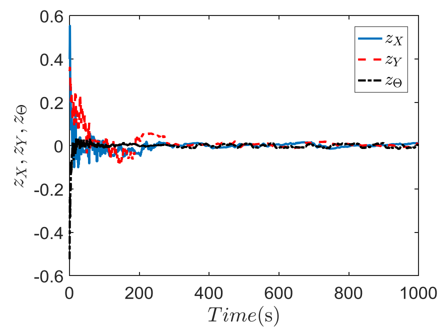

Figure 3. Position tracking errors.

It is shown in Fig. 2 that the NN weights fast converge to the suboptimal values during the online learning and control. Fig. 3 shows the position tracking errors z_X, z_Y, z_Θ and Figs. 4-6 show the centered position q of robot on the global Cartesian coordinate system. It can be seen that the tracking trajectory between the system output and the desired one satisfies the desired control performance.

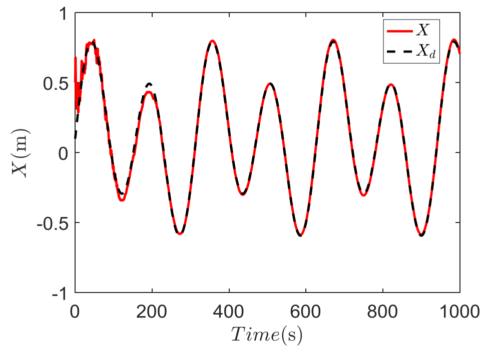

Figure 4. Evolution of X-position.

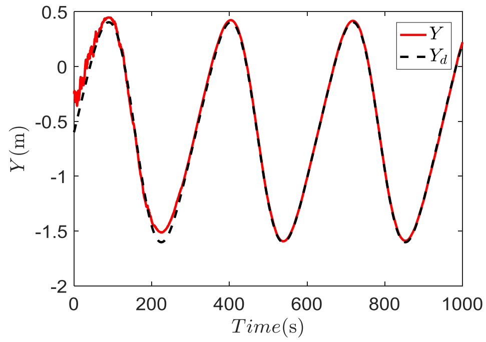

Figure 5. Evolution of Y-position.

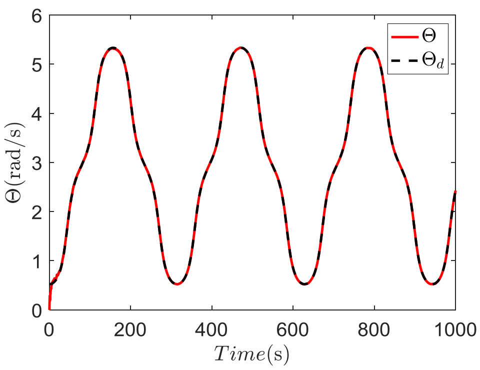

Figure 6. Evolution of Θ -position.

The triggering error and threshold are presented in Fig. 7. It is shown that the threshold is kept constant during inter-event interval, while the triggering error grow beneath the threshold all the time so that when it passes over the threshold, it is reset to zero. It can be also seen that the high triggering frequency occurs at the earlier stages of the control process but gradually reduces to the lower over time (see subplots in Fig. 7). Obviously, the sampled virtual velocities for the feedforward control laws, depicted in Figs. 8, 9, are held during inter-event intervals. Therefore, it can be concluded that the burden of the computational cost is mitigated.

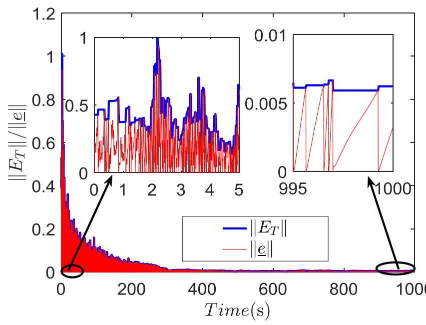

Figure 7. Evolution of sampling error and triggering threshold.

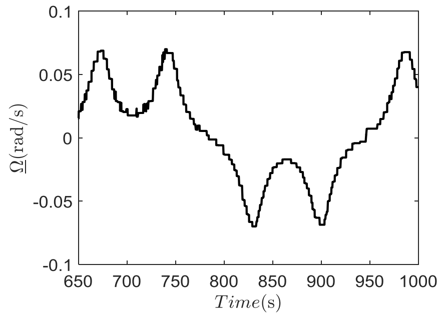

Figure 8. Evolution of sampled rotational velocity.

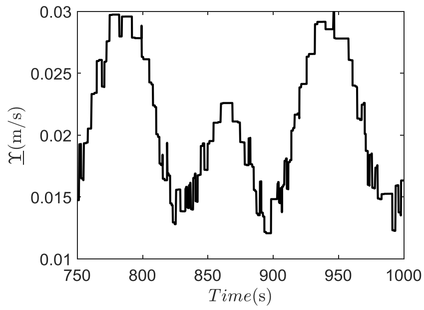

Figure 9. Evolution of sampled linear velocity.

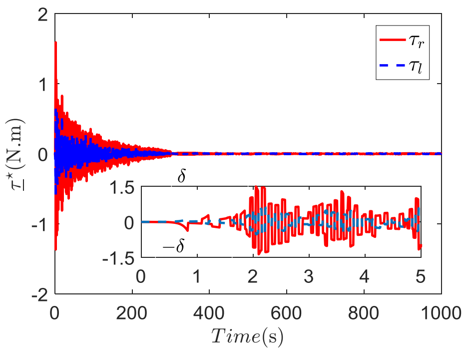

Figure 10. Robust optimal torques with dead-zone and saturation.

The even-triggered robust optimal torques with dead-zone and saturation is shown in Fig. 10, where the subplot shows the saturated left and right torques when they reach the maximum and minimum saturation limits δ and $-\delta$. During the online evolution, the torques decrease to the suboptimal values to make the value function (Fig. 11) be bounded and converged to the $\mathcal{L}_2$-

gain optimal value. Then, the solution of the even-triggering HJI equation is solved.

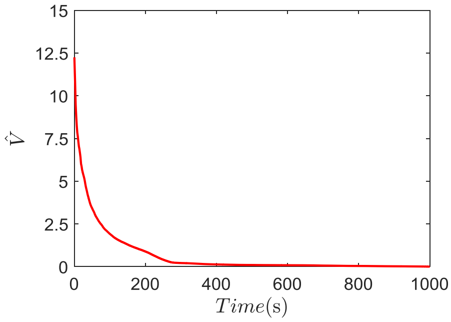

Figure 11. Convergence of the approximate optimal value function.

V. Conclusion

This paper provided a novel event-triggered robust optimal control method for mobile Euler-Lagrange systems in the presence of dead-zone and saturation actuators. The event-triggered feedforward control law was designed to transform the control problem with separate kinematics and dynamics into the equivalently integrated control problem. Based on the RL technique and the three-player differential game theory, the feedback robust optimal control scheme was then designed. The scheme included the event-triggered dead-zone disturbance law and the event-triggered constrained control law, where the parameters were only updated when a triggering condition held. Thanks to using RL, the proposed method relaxed the system identification procedure. The event-triggering condition was designed so that not only the Zeno phenomenon was excluded but also the closed system was guaranteed the stability while the cost function converged to the $\mathcal{L}_2$ -gain optimal. Future work will be concentrated on designing event-triggered controllers for completely unknown systems or multi-agent systems in consensus.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

Tan-Luy Nguyen conducted the theory analysis and simulation. Huu-Toan Tran, Trong-Toan Tran and Cong-Thanh Pham commented and revised the paper; all authors had approved the final version.

Acknowledgment

The authors would like to thank the Vietnam National Foundation for Science and Technology Development (NAFOSTED) for the support under the grant number 107.04-2019.25.

References

[1] L. N. Tan, “Omnidirectional vision-based distributed optimal tracking control for mobile multi-robot systems with kinematic and dynamic disturbance rejection,” *IEEE Trans. Ind. Electron.*, vol. 65, no. 7, pp. 5693–5703, Dec. 2018.

[2] R. Fierro and F. L. Lewis, “Control of a nonholonomic mobile robot using neural networks,” *IEEE Trans. Neural Netw*., vol. 9, no. 4, pp. 589–600, July 1998.

[3] N. Ahmed, J. Cortes, and S. Martinez, “Distributed control and estimation of robotic vehicle networks: Overview of the special issue,” *IEEE Control Syst. Mag*., vol. 36, no. 2, pp. 36–40, Mar. 2016.

[4] A. Abdessameud, A. Tayebi, and I. G. Polushin, “Leader-follower synchronization of Euler-Lagrange systems with time-varying leader trajectory and constrained discrete-time communication,” *IEEE Trans. Autom. Control*, vol. 62, no. 5, pp. 2539–2545, Aug. 2017.

[5] Q. Zhou, S. Zhao, H. Li, R. Lu, and C. Wu, “Adaptive neural network tracking control for robotic manipulators with dead zone,” *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 30, no. 12, pp. 3611–3620, Oct. 2019.

[6] R. R. Selmic and F. L. Lewis, “Dead-zone compensation in motion control systems using neural networks,” *IEEE Trans. Autom. Control*, vol. 45, no. 4, pp. 602–613, Apr. 2000.

[7] A. H. Tahoun, “Anti-windup adaptive PID control design for a class of uncertain chaotic systems with input saturation,” *ISA Transactions*, 66, pp. 176 – 184, 2017.

[8] A. H. Tahoun, “A new online delay estimation-based robust adaptive stabilizer for multi-input neutral systems with unknown actuator nonlinearities,” *ISA Transactions*, vol. 70, pp. 139–148, 2017.

[9] A. H. Tahoun, “Time-varying multiplicative/additive faults compensation in both actuators and sensors simultaneously for nonlinear systems via robust sliding mode control scheme,” *J. Franklin Insti.*, vol. 356, no. 1, pp. 103–128, 2019.

[10] Z. Liu, F. Wang, and Y. Zhang, “Adaptive visual tracking control for manipulator with actuator fuzzy dead-zone constraint and unmodeled dynamic,” *IEEE Trans. Syst. Man Cybern. Syst*., vol. 45, no. 10, pp. 1301–1312, Apr. 2015.

[11] W. He, A. O. David, Z. Yin, and C. Sun, “Neural network control of a robotic manipulator with input deadzone and output constraint,” *IEEE Trans. Syst. Man Cybern. Syst*., vol. 46, no. 6, pp. 759–770, Aug. 2016.

[12] L. N. Tan, “Event-triggered distributed H∞ control of physically interconnected mobile Euler–Lagrange systems with slipping, skidding and dead zone,” *IET Control Theory & Appl*., vol. 14, no. 3, Dec. 2019.

[13] Y. Zhao, Z. Duan, and G. Wen, “Finite-time consensus for secondorder multi-agent systems with saturated control protocols,” *IET Control Theory Appl*., vol. 9, no. 3, pp. 312–319, Feb. 2015.

[14] H. Su, M. Z. Q. Chen, X. Wang, and J. Lam, “Semiglobal observerbased leader-following consensus with input saturation,” *IEEE Trans. Ind. Electron*., vol. 61, no. 6, pp. 2842–2850, Aug. 2014.

[15] M. Cai and Z. Xiang, “Adaptive finite-time consensus tracking for multiple uncertain mechanical systems with input saturation,” *Int. J. Robust Nonlinear Control*, vol. 27, no. 9, pp. 1653–1676, 2017.

[16] Q. Zhou, H. Li, C. Wu, L. Wang, and C. K. Ahn, “Adaptive fuzzy control of nonlinear systems with unmodeled dynamics and input saturation using small-gain approach,” *IEEE Trans. Syst., Man, Cybern., Syst*., vol. 47, no. 8, pp. 1979–1989, Aug. 2017.

[17] Z. Chen, Z. Li, and C. L. P. Chen, “Adaptive neural control of uncertain MIMO nonlinear systems with state and input constraints,” *IEEE Trans. Neural Netw. Learn. Syst*., vol. 28, no. 6, pp. 1318–1330, Mar. 2017.

[18] L. N. Tan, “Distributed H_∞ optimal tracking control for strict-feedback nonlinear large-scale systems with disturbances and saturating actuators,” *IEEE Trans. Syst., Man, Cybern., Syst.*, (to be published, 2018.

[19] R. S. Sutton and A. G. Barto, *Reinforcement Learning-an Introduction*, Cambridge, MA: MIT Press, 1998.

[20] T. Dierks, B. Brenner, and S. Jagannathan, “Neural network-based optimal control of mobile robot formations with reduced information exchange,” *IEEE Trans. Control Syst. Technol.*, vol. 21, no. 4, pp. 1407–1415, July 2013.

[21] C. Mu and K. Wang, “Approximate-optimal control algorithm for constrained zerosum differential games through event-triggering mechanism,” *Nonlinear Dyn.*, vol. 95, no. 4, pp. 2639–2657, 2019.

[22] F. Tatari, K. G. Vamvoudakis, and M. Mazouchi, "Optimal distributed learning for disturbance rejection in networked non-linear games under unknown dynamics," *IET Control Theory Appl*., vol 13, no. 17, pp. 2838 – 2848, Nov. 2019.
[23] Y. Huang, "Optimal guaranteed cost control of uncertain non-linear systems using adaptive dynamic programming with concurrent learning," *IET Control Theory Appl*., vol. 12, no. 8, pp. 1025–1035, May 2018.
[24] P. Tabuada, "Event-triggered real-time scheduling of stabilizing control tasks," *IEEE Trans. Autom. Control*, vol. 52, no. 9, pp. 1680–1685, Apr. 2007.
[25] Z. Sun, L. Dai, Y. Xia and K. Liu, "Event-based model predictive tracking control of nonholonomic systems with coupled input constraint and bounded disturbances," *IEEE Trans. Autom. Control*, vol. 63, no. 2, pp. 608–615, Aug. 2018.
[26] V. Narayanan, S. Jagannathan, "Event-triggered distributed approximate optimal state and output control of affine nonlinear interconnected systems," *IEEE Trans. Neural Netw. Learn. Syst*., vol. 29, no. 7, pp. 2846–2856, June 2018.
[27] L. Ding, Q. Han, X. Ge, and X. Zhang, "An overview of recent advances in eventtriggered consensus of multiagent systems," *IEEE Trans. Cybern.*, vol. 48, no. 4, pp. 1110–1123, Nov. 2018.
[28] H. K. Khalil, *Nonlinear Systems*, 3rd ed. Englewood Cliffs, NJ, USA: PrenticeHall, 2002.
[29] P. Ioannou and J. Sun, *Robust Adaptive Control*, New Jersey: Prentice Hall, 1996.
[30] Y. Tan, R. Dong, and R. Li, "Recursive identification of sandwich systems with dead zone and application," *IEEE Trans. Control. Syst. Technol.*, vol. 17, no. 4, pp. 945–951, May 2009.
[31] N. T. Luy, N. T. Thanh, and H. M. Tri, "Reinforcement learning-based intelligent tracking control for wheeled mobile robot," *Transactions of the Institute of Measurement and Control*, vol. 36, no. 7, pp. 868–877, 2014.
[32] C. Wen, J. Zhou, Z. Liu, and H. Su, "Robust adaptive control of uncertain nonlinear systems in the presence of input saturation and external disturbance," *IEEE Trans. Autom. Control*, vol. 56, no. 7, pp. 1672–1678, Mar. 2011.
[33] P. Tabuada, "Event-triggered real-time scheduling of stabilizing control tasks," *IEEE Trans. Autom. Control*, vol. 52, no. 9, pp. 1680–1685, Sep. 2007.
[34] L. N. Tan, "Event-triggered distributed H∞ constrained control of physically interconnected large-scale partially unknown strict-feedback systems," *IEEE Trans. Syst., Man, and Cybern. Syst.* (to be published, 2019.
[35] M. Abu Khalaf and 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, 2005.
[36] T. Basar and P. Bernhard, *Optimal Control and Related Minimax Design Problems: A Dynamic Game Approach*, MA: Birkhuser, 1995.
[37] D. Vrabie and F. Lewis, "Neural network approach to continuous-time direct adaptive optimal control for partially unknown nonlinear systems," *Neural Netw*., vol. 22, no. 3, pp. 237–246, 2009.

Tan-Luy Nguyen received the M.Sc. degree in automation and control engineering and the Ph.D. degree in Automation from the Ho Chi Minh City University of Technology, Ho Chi Minh City, Vietnam, in 2006 and 2015, respectively.

He is currently with the Faculty of Electronics Technology, Industrial University of Ho Chi Minh City, Vietnam. His current research interests include adaptive dynamic programming, distributed control, large-scale systems, and deep learning.

Huu-Toan Tran received the M.Sc. degrees in automation from HCM City University of Technology, Vietnam in 2009, and the Ph.D. degree in electronics science and technology from University of Electronics Science and Technology of China in 2015.

He is currently a Lecturer in the Faculty of Electronics Technology, Industrial University of Ho Chi Minh City, Vietnam. His current research interests include control application, robot learning, and human-robot interaction.

Trong-Toan Tran received his B.Eng. and M.Eng. degrees both in automation from Bauman Moscow State Technical University, Moscow, Russia, in 2006 and 2008, respectively, and the Ph.D. degree in automatic control from University of Electronic Science and Technology of China, China, in 2016.

He was a Research Fellow with the Department of Electrical and Computer Engineering, National University of Singaport (NUS), from 2017 to 2018. He is currently a research scientist in control and mechatronics Engineering at the National Institute of Applied Mechanics and Informatics, Vietnam Academy of Science and Technology. His current research interests include autonomous control, unmanned aerial vehicle, deep-water technology, and industrial robotic system.

Cong-Thanh Pham received the M.Sc. degree in automation and control engineering from the Ho Chi Minh City University of Technology, Ho Chi Minh City, Vietnam, 2006, the Ph.D. received degree at Department of Control Science and Engineering, Huazhong University of Science and Technology (HUST), Wuhan 430074, China, 2014.

He is currently with the Faculty of Automation Control, Ho Chi Minh City University of Food & Industrial, Ho Chi Minh City. His current research interests include AC motor, power electronics, adaptive dynamic programming.