Neural Networks 114 (2019) 78–90

Contents lists available at ScienceDirect

Neural Networks

journal homepage: www.elsevier.com/locate/neunet

Optimization of sampling intervals for tracking control of nonlinear systems: A game theoretic approach[☆]

Avimanyu Sahoo[a,*], Vignesh Narayanan[b]

[a] Division of Engineering Technology, Oklahoma State University, Stillwater, OK, USA
[b] Department of Electrical and Systems Engineering, Washington University, St. Louis, MO, USA

ARTICLE INFO

Article history:
Received 26 August 2018
Received in revised form 20 January 2019
Accepted 26 February 2019
Available online 8 March 2019

Keywords:
Approximate dynamic programming
Event-based control
Hamilton–Jacobi–Issac equation
Nonlinear tracking control
Optimal control
Zero-sum-game

ABSTRACT

This paper presents a near optimal adaptive event-based sampling scheme for tracking control of an affine nonlinear continuous-time system. A zero-sum game approach is proposed by introducing a novel performance index. The optimal value function, i.e., the solution to the associated Hamilton–Jacobi–Issac (HJI) equation is approximated using a functional link neural network (FLNN) with event-based aperiodic state feedback information as inputs. The saddle point approximated optimal solution is employed to design the near optimal event-based control policy and the sampling condition. An impulsive weight update scheme is designed to guarantee local ultimate boundedness of the closed-loop parameters, which is analyzed via extension of Lyapunov stability theory for the impulsive hybrid dynamical systems. Zeno-freeness of the event-sampling scheme is enforced and its effect on stability is analyzed. Finally, numerical simulation results are included to corroborate the analytical design, which shows a 48.82% reduction of feedback communication and computational load.

1. Introduction

Digital implementation of control schemes use traditional sampled-data based control approach (Astrom & Wittenmark, 2013) with a fixed sampling period, selected *a priori* based on worst-case scenario. Recently developed event-based sampling and control (ETC) (Cheng & Ugrinovskii, 2016; Heemels & Donkers, 2013; Peng, Song, Xie, Zhao, & Fei, 2016; Postoyan et al., 2015; Tabuada, 2007; Tallapragada & Chopra, 2013) and self-triggered control (STC) (Gommans, Antunes, Donkers, Tabuada, & Heemels, 2014; Mazo, Anta, & Tabuada, 2010) schemes show that fixed sampling schemes lead to redundant use of computational and communication resources. Both ETC (Cheng & Ugrinovskii, 2016; Heemels & Donkers, 2013; Peng et al., 2016; Postoyan et al., 2015; Tabuada, 2007; Tallapragada & Chopra, 2013) and STC (Gommans et al., 2014; Mazo et al., 2010) schemes provide a unified stability and performance based approach for determining the sampling intervals and the frequency of the controller execution. Therefore, the ETC (Cheng & Ugrinovskii, 2016; Heemels & Donkers, 2013; Peng et al., 2016; Postoyan et al., 2015; Tabuada, 2007; Tallapragada & Chopra, 2013) and the STC (Gommans et al., 2014; Mazo et al., 2010) schemes are resource aware control approaches, which can save computational and communication resources more effectively. In the ETC paradigm (Heemels & Donkers, 2013; Tabuada, 2007; Tallapragada & Chopra, 2013), sampling instants and frequency of control execution are typically orchestrated via a sampling/triggering mechanism, whereas this extra piece of hardware for triggering mechanism is replaced by a software based sampling scheme in STC (Gommans et al., 2014; Mazo et al., 2010). The STC determines the next sampling instant using the current and the past feedback information. An ample amount of research results, both on ETC and STC schemes, are available in the literature (Gommans et al., 2014; Heemels & Donkers, 2013; Mazo et al., 2010; Tabuada, 2007; Tallapragada & Chopra, 2013) for regulation as well as tracking problems.

Trajectory tracking control (Cheng & Ugrinovskii, 2016; Gao & Chen, 2008; Peng et al., 2016; Postoyan et al., 2015; Tallapragada & Chopra, 2013), where the system state or the output tracks a desired trajectory, has a wide range of practical applications. Leader follower architecture of mobile robots (Cheng & Ugrinovskii, 2016; Han, Lu, & Chen, 2015), autonomous systems employed for surveillance (Postoyan et al., 2015), target tracking radars are few of the examples. In a trajectory tracking control, the control policy is, in general, time varying which leads to a continuous control effort to keep the tracking error minimum. Therefore, a time-based periodic implementation of a tracking control policy requires significantly higher computations when

☆ This research is funded by start-up fund Oklahoma State University, USA.
* Corresponding author.
E-mail addresses: avimanyu.sahoo@okstate.edu (A. Sahoo), vnxv4@mst.edu (V. Narayanan).

https://doi.org/10.1016/j.neunet.2019.02.008
0893-6080/

compared to a state regulation problem. Event-based tracking control (Cheng & Ugrinovskii, 2016; Postoyan et al., 2015; Tallapragada & Chopra, 2013) is most suitable in this scenario as it can potentially reduce the computations without jeopardizing the tracking performance.

In ETC and STC formalism, several tracking control schemes are presented by the authors Cheng and Ugrinovskii (2016), Tallapragada and Chopra (2013), Peng et al. (2016) and Postoyan et al. (2015) and the references there in. In Tallapragada and Chopra (2013), the authors presented an emulation based event-triggered controller and event-triggering condition design for state tracking problem, which guarantees ultimate boundedness of the tracking error. An output tracking control for system with communication constraints such as delays and packet losses is presented in Peng et al. (2016). The authors proposed an improved triggering mechanism to guarantee H_∞ tracking performance. Further, experimental results on reduction in communication are presented in Postoyan et al. (2015) for unicycle mobile robot tracking control using event-based approach. Leader follower tracking control of multi-agent system is proposed in Cheng and Ugrinovskii (2016) for both directed and undirected graphs, where the authors guaranteed the boundedness of tracking errors via Lyapunov stability analysis. The governing design principle in all the above event-based tracking control schemes (Cheng & Ugrinovskii, 2016; Peng et al., 2016; Postoyan et al., 2015; Tallapragada & Chopra, 2013) is to retain the stability of the system by designing a triggering condition while implementing the continuous control policy in an event-based frame work.

Optimal control (Lewis, Vrabie, & Syrmos, 2012) approaches are used to optimize the cumulative performance cost of a system. In general, the Bellman equation (Bellman, 2013) and the Hamilton–Jacobi–Bellman (HJB) equation (Bertsekas, Bertsekas, Bertsekas, & Bertsekas, 1995) are employed to obtain the optimal solutions, respectively for discrete and continuous time systems. A closed form analytical solution to the HJB partial differential equation is almost impossible to obtain (Bellman, 2013; Sutton & Barto, 1998). In this case, adaptive dynamic programming (ADP) and reinforcement learning (RL) (Bertsekas et al., 1995; Sutton & Barto, 1998) based approaches are employed to approximate the solution of the HJB equation. Invoking the universal approximation property (Lewis, Jagannathan, & Yesildirak, 1998), ADP/RL schemes use neural networks (NN) for obtaining an approximate solution to the HJB equation. Adaptive critic framework, commonly referred to as actor-critic in ADP/RL literature, is widely used to solve the optimal control problem where the critic NN is used to approximate the optimal value function and the actor NN is employed to approximate the optimal control policy.

A great deal of research results under ADP/RL schemes, such as, value iteration (Bertsekas et al., 1995), policy iteration (Bertsekas, 2017), online learning (Dierks & Jagannathan, 2009), robust ADP (Wang, Liu, Zhang, & Li, 2018), and data driven ADP (Zhang, Cui, Zhang, & Luo, 2011) are available in the literature for both optimal regulation (Heydari, 0000) and tracking (Heydari & Balakrishnan, 2014; Kamalapurkar, Dinh, Bhasin, & Dixon, 2015; Modares & Lewis, 2014; Mu, Sun, Song, & Yu, 2016; Wang et al., 2018; Yang & He, 2018; Zhang et al., 2011) control problems. For the tracking control problem, the authors in Mu et al. (2016) presented an iterative globalized dual heuristic programming (GDHP) based solution for discrete time systems. A time-based non-iterative approach, using the time history for the utility function and the activation function, is presented by the authors in Dierks and Jagannathan (2009), where the optimal solution is obtained online and forward in time. A self learning robust ADP/RL approach is presented in Yang and He (2018) in the presence of mismatched disturbance. Further, the authors in Modares and Lewis (2014) presented a constrained input ADP/RL design for partially-unknown systems using integral reinforcement learning. In Kamalapurkar et al. (2015) and Zhang, Wei, and Luo (2008) the ADP/RL based tracking control is designed using an augmented system which transformed the time-varying value function to a time-invariant form. All the above designs use an infinite horizon performance index. For the finite time optimal control problem, an ADP/RL based optimal tracking control is presented by the authors in Heydari and Balakrishnan (2014) for a continuous time system. However, all the above designs (Heydari & Balakrishnan, 2014; Kamalapurkar et al., 2015; Modares & Lewis, 2014; Mu et al., 2016; Yang & He, 2018; Zhang et al., 2011) use continuous feedback information for NN based approximation and controller execution. Thus, the designs are computationally intensive.

Event-based optimal control designs using ADP/RL can reduce the computational burden on the embedded processors and are studied by the authors in Liu, Tang, and Fei (2016), Narayanan and Jagannathan (2016), Sahoo, Xu, and Jagannathan (2017), Vamvoudakis, Mojoodi, and Ferraz (2017), Wang, Mu, Zhang, and Liu (2016), Wang and Liu (0000), Zhang, Zhao, and Zhu (2017) and Zhong and He (2017) and the references there in. The authors in Sahoo et al. (2017) presented optimal state regulation using event-based ADP/RL for an isolated nonlinear continuous time system. The authors utilized the event-based states to update the NN weights and obtained the solution of the HJB equation. An adaptive triggering condition is designed utilizing the Lyapunov stability analysis. The event-based ADP scheme is extended for interconnected system in Narayanan and Jagannathan (2016). An observer-based ADP/RL scheme in the context of event-based sampling is presented in Zhong and He (2017) where a NN based observer is employed to reconstruct the system states and the NN weights are tuned only at the triggering instants. The triggering condition is derived using the Lyapunov stability analysis. From the tracking control perspective, the authors in Vamvoudakis et al. (2017) proposed an actor-critic based ADP scheme with a discounted cost function. The event-based sampling instants are determined such that the approximated optimal control policy retains the stability of the system.

The optimal control problem in an event-based framework is twofold: (1) optimization of the control policy; and (2) optimization of sampling instants. However, all the above event-based ADP/RL near optimal designs (Liu et al., 2016; Narayanan & Jagannathan, 2016; Sahoo et al., 2017; Vamvoudakis & Ferraz, 2016; Vamvoudakis et al., 2017; Wang & Liu, 0000; Zhong & He, 2017) optimized the control policy such that it can minimize the corresponding performance index whereas the triggering condition is designed to retain the stability and, hence, not optimal. Since, the design of the triggering condition affects the optimal cost (Vamvoudakis et al., 2017) and the approximation accuracy (Sahoo et al., 2017), a trade-off must be achieved. In our recent work, Sahoo, Narayanan, and Jagannathan (2017), a unified design of sampling intervals and control policy is presented to regulate linear systems using the zero-sum game (Başar & Bernhard, 1995). However, optimization of the sampling interval to reach a trade-off between the optimal cost and the approximation accuracy for tracking control of nonlinear systems is still a challenging and open problem.

Motivated by the above limitations, in this paper, a performance based sampling scheme, i.e., an improved triggering condition, is presented for trajectory tracking control. A novel performance index is introduced which aids in designing the optimal triggering threshold. The event-based tracking error system is formulated by incorporating the error between the continuous control policy and the event-based control policy, referred to as *sampled error policy*. An augmented tracking error system, similar to Kamalapurkar et al. (2015) and Zhang et al. (2008), but, in an event-based formalism, is defined. The novel time-invariant performance index translates the event-based optimization problem to a min–max problem and the saddle point solution to the zero-sum game (Başar & Bernhard, 1995) is obtained. Thus, the designed control policy minimizes the performance index in the presence of the worst case error in the control policy due to aperiodic sampling.

The solution to the Hamilton–Jacobi–Isac (HJI) equation, i.e., the optimal value function, is approximated using a functional link critic neural network. Event-based intermittently available state information is used as the input of the NN. The estimated optimal control policy and the triggering threshold are computed from the approximated value function. Since, the triggering threshold is a function of the estimated NN weights, which are updated with an impulsive update scheme, the triggering condition becomes adaptive. To avoid Zeno behavior of the closed-loop system during the learning period, a lower bound on the adaptive threshold is enforced. The ultimate boundedness of the tracking and the NN weight estimation errors are proved using the extension of the Lyapunov theorem (Haddad, Chellaboina, & Nersesov, 0000) for hybrid dynamical systems.

Contribution. The main contributions of the paper are: (1) definition of a novel performance function to simultaneously optimize the control policy and determine the threshold for the event-triggering condition for tracking control problem; (2) impulsive on-line NN weight update scheme using the state and input information only at the triggering instants; (3) design of novel, adaptive event-triggering condition with Zeno-free behavior; and (4) proof of stability using extension for impulsive hybrid system.

Notations. The notations used in this paper are standard. The n-dimensional Euclidean space is denoted by $\Re^n$. The set $\mathbb{N}$ is the set of natural numbers. The operator Δ is the difference operator, ∇ is the gradient operator, $(\cdot)^T$ is the transpose of a vector or matrix $(\cdot)$, λ_{min}, and λ_{max} are the minimum and maximum eigenvalues. The norm operator $\|\cdot\|$ is the Euclidean norm for vectors and Frobenius norm for matrices. The function argument t is dropped for brevity and made explicit only when it is necessary. For example, $x(t)$ is represented as x.

The remainder of the paper is organized as follows. Section 2 formulates the problem considered in this paper. In Section 3 a novel performance index is proposed along with the optimization scheme. An approximate solution using ADP is presented in Section 4 followed by the simulation results in Section 5 and conclusions in Section 6. The Appendix presents the proof of all the lemmas and theorems.

2. Background and problem definition

2.1. Background

Consider an affine nonlinear continuous-time system given by

$$\dot{x}(t) = f(x(t)) + g(x(t))u(t), \quad x(0) = x_0 \tag{1}$$

where $x : [0, \infty) \to \Re^n$ and $u : \Re^n \to \Re^m$ are state and control input vectors, respectively. The nonlinear vector function $f : \Re^n \to \Re^n$ and matrix function $g : \Re^n \to \Re^{n\times m}$ are, respectively, the internal dynamics and the input gain function with $f(\mathbf{0}) = \mathbf{0}$ where $\mathbf{0}$ is a vector of all zeros of appropriate dimension.

The control objective is to track a feasible reference trajectory $x_d(t) \in \Re^n$ generated by a reference system represented by

$$\dot{x}_d(t) = \zeta(x_d(t)), \quad x_d(0) = x_{d0} \tag{2}$$

where $x_d(t) \in \Re^n$ is the reference state, and $\zeta : \Re^n \to \Re^n$ is the internal dynamics with $\zeta(\mathbf{0}) = \mathbf{0}$. Following standard characteristics of the systems in (1) and (2) are assumed.

Assumption 1. System (1) is stabilizable and the system states are available for measurement.

Assumption 2. The functions $f(x)$ and $g(x)$ are Lipschitz continuous for all $x \in \Omega_x$ where Ω_x is a compact set containing the origin. Further, the function $g(x)$ has a full column rank for all $x \in \Omega_x$ and satisfies $\|g(x)\| \le g_M$ for some constant $g_M > 0$. In addition, $g(x_d)g^+(x_d) = I$ where $g^+ = (g^Tg)^{-1}g^T$.

Assumption 3. The feasible reference trajectory $x_d(t) \in \Omega_{x_d}$, where Ω_{x_d} is a compact set, is bounded such that $\|x_d(t)\| \le b_{x_d}$ where $b_{x_d} > 0$ is a constant.

Define the error between the system state and the reference state as tracking error, given by $e_r(t) \triangleq x(t) - x_d(t)$. Then, the tracking error system, utilizing (1) and (2), can be defined by

$$\dot{e}_r = \dot{x} - \dot{x}_d = f(e_r + x_d) + g(e_r + x_d)u - \zeta(x_d). \tag{3}$$

The steady state feed forward control policy for the reference trajectory (Kamalapurkar et al., 2015) can be expressed as

$$u_d = g^+(x_d)(\zeta(x_d)) - f(x_d) \tag{4}$$

where $u_d : \Re^n \to \Re^m$ is the steady state control policy corresponding to the reference trajectory. By augmenting the tracking error e_r and desired trajectory x_d, the dynamics of the augmented tracking error system can be represented as

$$\dot{\chi} = F(\chi) + G(\chi)w \tag{5}$$

where $\chi \triangleq [e_r^T \quad x_d^T]^T \in \Re^{2n}$ is the augmented state with $\chi(0) = [e_r^T(0) \quad x_d^T(0)]^T = \chi_0$, $F : \Re^{2n} \to \Re^{2n}$ is given by $F(\chi) \triangleq \begin{bmatrix} f(e_r + x_d) + g(e_r + x_d)u_d - \zeta(x_d) \\ \zeta(x_d) \end{bmatrix}$, $G : \Re^{2n} \to \Re^{2n\times m}$ given by $G(\chi) \triangleq \begin{bmatrix} g(e_r + x_d) \\ \mathbf{0} \end{bmatrix}$, and the mismatched control policy $w \triangleq u - u_d \in \Re^m$. It is routine to check that $F(\mathbf{0}) = \mathbf{0}$.

The infinite horizon performance index with state constraint enforced by the dynamical system in (5) can be defined as

$$J(\chi, w) = \int_0^\infty [\chi^T(\tau)\bar{Q}\chi(\tau) + w(\tau)^T Rw(\tau)]d\tau \tag{6}$$

where $\bar{Q} \triangleq \begin{bmatrix} Q & \mathbf{0}_{n\times n} \\ \mathbf{0}_{n\times n} & \mathbf{0}_{n\times n} \end{bmatrix} \in \Re^{2n\times 2n}$ with $Q \in \Re^{n\times n}$ and $R \in \Re^{m\times m}$ are symmetric positive definite matrices. The matrix $\mathbf{0}_{n\times n}$ is a matrix with all elements zero. Note that the performance index is defined using the mismatched policy w (Kamalapurkar et al., 2015) and, therefore, the cost functional is finite for any admissible control policy $w \in \Omega_w$ where Ω_w is the set of all admissible policies (Lewis et al., 2012).

In traditional ADP/RL approach, an optimal control policy w can be designed by approximating the solution of the HJB equation using NNs (Kamalapurkar et al., 2015) with the continuously available state information as input. However, in ETC paradigm, the feedback information are available intermittently only at the triggering instants. This aperiodic availability of the state information introduces an error in the system dynamics (Tabuada, 2007), and, further, influences the approximation accuracy (Sahoo et al., 2017) and the optimal value of cost (Vamvoudakis & Ferraz, 2016). Therefore, to achieve a trade-off, the optimization problem in ETC context requires co-designing of the sampling intervals and the control policy. With this effect, the event-based optimal tracking control problem, in hand, is formally defined next.

2.2. Problem definition

In an ETC formalism, define the aperiodic event-based sampling instants as a sequence $\{t_k\}_{k\in\{0,\mathbb{N}\}} \subset \{t\}$, with $t_0 = 0$ and $t_k < t_{k+1}, \forall k \in \{0, \mathbb{N}\}$. The system state $x(t_k)$ is sent to the controller at the time instants $t_k, \forall k \in \{0, \mathbb{N}\}$. The sampled state at the controller can be written as

$$x_s(t) = x(t_k), \ \forall t \in [t_k, t_{k+1}). \tag{7}$$

The event-based control policy, executed at the sampling instants $t_k, \forall k \in \{0, \mathbb{N}\}$ with the state information $x_s(t)$ and reference trajectory $x_{ds} = x_d(t_k), \forall t \in [t_k, t_{k+1})$, can be represented as

$$u_s(t) = \mu(x_s(t), x_{ds}(t)) = \mu(e_r(t), t_k), \forall t \in [t_k, t_{k+1}) \tag{8}$$

where $\mu : \Re^n \times \Re_{\geq 0} \to \Re^m$. The control policy $u_s(t)$ is held at the actuator using a zero-order hold and applied to the system till the next update. Therefore, the event-based control policy is a piecewise constant function.

The event-based nonlinear system can be rewritten from (1) with event-based control policy u_s as

$$\dot{x} = f(x) + g(x)u_s. \tag{9}$$

We define the error between the continuous control input $u(t) \triangleq \mu(x(t), x_d(t)) = \mu(e_r(t), t_k)$ and the event-based control input $u_s(t)$ in (8) as sampled error policy, $e_u(t) \in \Re^m$, which is given by

$$e_u(t) = u_s(t) - u(t). \tag{10}$$

The event-triggered system dynamics in (9), with sampled error policy in (10), lead to

$$\dot{x} = f(x) + g(x)u + g(x)e_u. \tag{11}$$

The tracking error dynamics, with the event-triggered system (11) and the reference system (2), become

$$\dot{e}_r = f(e_r + x_d) + g(e_r + x_d)u + g(e_r + x_d)e_u - \zeta(x_d). \tag{12}$$

The event-based augmented tracking error system, from (12) and (2), can be represented as

$$\dot{\chi} = F(\chi) + G(\chi)w + G(\chi)e_u. \tag{13}$$

Our main objective is to develop a unified design scheme by minimizing a performance index such that both the sampling intervals and the control policy are optimized. Therefore the problem in hand is threefold: (1) redefinition of the time-invariant performance index in (6) to obtain the worst case threshold for the sampled error policy, e_u in (13), which in turn determines the event-based sampling intervals; (2) design of the event-based sampling condition such that the sampling intervals are maximized; and (3) design of the NN weight update law for approximation of the solution of the corresponding HJI equation in an ETC framework. Solution to the above problem is presented in the next section.

3. Novel performance index and solution to the min–max optimization problem

In this section, the co-design problem is formulated as a min–max optimization problem by introducing a novel performance index. A saddle point solution to the optimization problem is obtained by solving the associated HJI equation.

3.1. Reformulation of performance index

The event-based sampling instants can be designed by triggering the events when the sampled error policy, e_u, in (10) reaches a maximum value without jeopardizing the stability. Alternatively, the design is to obtain an optimal threshold for $e_u(t)$, to determine the triggering instants, with respect to the desired performance of the system. With this effect, redefine the cost functional (6) with the dynamic constraint (13) as

$$J(\chi, w, \hat{e}_u) = \int_0^\infty [\chi^T(\tau)\bar{Q}\chi(\tau) + w(\tau)^T Rw(\tau) - \gamma^2 \hat{e}_u^T(\tau)\hat{e}_u(\tau)]d\tau \tag{14}$$

where $\gamma > \gamma^*$ (Başar & Bernhard, 1995) represents the penalizing factor for the threshold $\hat{e}_u$ of sampled error policy e_u with $\gamma^* > 0$ such that the performance index is finite for all admissible w.

Note that the performance index (14) is minimized by maximizing the threshold $\hat{e}_u$ and minimizing the control policy w. Therefore, the optimization problem leads to a min–max problem having two players. The mismatch control policy, w, acts as player 1, i.e., the minimizing player, and the threshold, $\hat{e}_u$, as player 2, i.e., the maximizing player. The objective, now, can be redefined as solving the two player zero-sum-game (Başar & Bernhard, 1995) to reach at the saddle-point optimal value, $V^* : \Re^{2n} \to \Re_{\geq 0}$, i.e., the optimal value where $min_w \ max_{\hat{e}_u} \ J(\chi, w, \hat{e}_u) = max_{\hat{e}_u} \ min_w \ J(\chi, w, \hat{e}_u)$.

Remark 1. The traditional min–max problem (Başar & Bernhard, 1995) optimizes the control policy in the presence of an independent disturbance, explicitly added to the system dynamics, by using a performance index similar to (14). However, the threshold $\hat{e}_u$ in (14) is implied in the system dynamics (12). The main advantage of the proposed performance index is to obtain the worst case value of $\hat{e}_u$ in terms of system state χ, which can be used as threshold for sampled error policy e_u.

From (14) the saddle-point optimal value, V^*, can be written as

$$V^*(\chi) = \min_{w(\tau)|\tau\in\Re_{\geq t}} \ \max_{\hat{e}_u(\tau)|\tau\in\Re_{\geq t}} \int_t^\infty \Big[\chi^T(\tau)\bar{Q}\chi(\tau) + w(\tau)^T Rw(\tau) - \gamma^2 \hat{e}_u^T(\tau)\hat{e}_u(\tau)\Big] d\tau. \tag{15}$$

Define the Hamiltonian, with the admissible control policy and dynamic constraint (13), as

$$\mathcal{H}(\chi, w, \hat{e}_u) = \chi^T\bar{Q}\chi + w^T Rw - \gamma^2 \hat{e}_u^T\hat{e}_u + V_\chi^{*T}[F + Gw + Ge_u] \tag{16}$$

where $V_\chi^* = \partial V^*/\partial\chi$.

By using the stationarity conditions, $\frac{\partial\mathcal{H}(\chi,w,\hat{e}_u)}{\partial w} = 0$ and $\frac{\partial\mathcal{H}(\chi,w,\hat{e}_u)}{\partial\hat{e}_u} = 0$, the optimal mismatch control policy

$$w^*(\chi) = -\frac{1}{2}R^{-1}G^T(\chi)V_\chi^*(\chi) \tag{17}$$

and the worst-case threshold value

$$e_u^*(\chi) = \frac{1}{2\gamma^2}G^T(\chi)V_\chi^*(\chi). \tag{18}$$

From (17) and (4) the continuous optimal control policy u^* is given by

$$u^* = -\frac{1}{2}R^{-1}G^T(\chi)V_\chi^*(\chi) + g^+(x_d)(\zeta(x_d) - f(x_d)). \tag{19}$$

To implement the control policy (19) in the ETC framework, the augmented state at the controller can be expressed as

$$\chi_s(t) = \chi(t_k), \ \ \forall t \in [t_k, t_{k+1}) \tag{20}$$

where $\chi(t_k) = [e_r^T(t_k), x_d^T(t_k)]^T, \forall k \in \{0, \mathbb{N}\}$. The event-sampled optimal control policy u_s^* with the augmented sampled state can be expressed as

$$u_s^* = -\frac{1}{2}R^{-1}G^T(\chi_s)V_{\chi_s}^* + g^+(x_{ds})(\zeta(x_{ds}) - f(x_{ds})). \tag{21}$$

where $V_{\chi_s}^* = \frac{\partial V^*(\chi)}{\partial \chi}|_{\chi=\chi_s}$.

The HJI equation with optimal policies (17) and (18) is given by

$$\mathcal{H}^* = \chi^T\bar{Q}\chi + w^{*T}Rw^* - \gamma^2 e_u^{*T}e_u^* + V_\chi^{*T}[F + Gw^* + Ge_u^*] = 0 \tag{22}$$

for all χ with $V^*(0) = 0$. The following assumption for the solution of the HJI equation is required to proceed.

Assumption 4. The solution to the HJI equation, i.e., the optimal value function V^*, exists and continuously differentiable.

Remark 2. Note that the solution of the HJI equation (22), i.e., the optimal value function, V^*, exists for a reachable and zero-state observable system for $\gamma > \gamma^*$, where γ^* is the H_∞ gain Başar and Bernhard (1995). Therefore, the assumption is trivial.

3.2. Optimal event-sampling condition and stability analysis

The sampling condition for the triggering mechanism can be defined using worst case sampled error e_u^* as threshold. Define the sampling condition as

$$t_{k+1} = \inf\{t > t_k | e_u(t)^T e_u(t) = \max\{r^2, \frac{1}{4\gamma^4}V_\chi^{*T}GG^TV_\chi^*\}\} \tag{23}$$

with $t_0 = 0$. The parameter $r > 0$ is a design choice.

Remark 3. The parameter $r > 0$ is introduced in the triggering condition to enforce the positive minimum inter-sample time, i.e., Zeno free behavior of the system and can be arbitrarily selected close to zero. However, this minimum threshold on triggering condition leads to a bounded stability of the ETC system and a detailed discussion is provided in Remark 4.

Next, the stability results of the optimal event-triggered system are presented in the theorem. The transformation of the time-varying tracking control problem to a time-invariant problem results in a positive semidefinite optimal value function (15) (Kamalapurkar et al., 2015), rendering it unsuitable to be treated as a Lyapunov candidate function. This short-coming is overcome by considering the time-variant counter part of the time-invariant optimal value function V^* given by $V_t^* : \Re^n \times \Re_{\geq 0} \to \Re$ where $V_t^*(e_r, t) = V^*([e_r^T, \ x_d^T(t)]^T)$ for all $e_r \in \Re^n$ and for all $t \in \Re_{\geq 0}$. It is shown in Kamalapurkar et al. (2015) that $V_t^* : \Re^n \times \Re_{\geq 0} \to \Re$ is a valid candidate Lyapunov function.

Before presenting the stability results, the following technical lemma, which also establishes the input-to-state stability (ISS) of the system (13) with respect to sampled error policy, is presented.

Lemma 1. *Consider the augmented tracking error system (13) and the performance index (14). Let the solution to the HJI equation (22) be V^* and Assumptions 1–4 hold. Then the mismatch optimal policy (17) renders the augmented system (13) local ISS with respect to $e_u(t)$.*

Proof. Refer to Appendix.

Theorem 1. *Consider the affine nonlinear system (1) and the reference system (2), reformulated as event-sampled augmented tracking error system (13). Let V^* be the solution of the HJI equation (22) and Assumptions 1–4 hold. Then, with the event-based optimal control policy (21) and event-based sampling condition (23), the tracking error is ultimately bounded provided γ satisfies $\gamma^2 > \lambda_{max}(R)$.*

Proof. Refer to Appendix.

Remark 4. From the proof of Theorem 1 for Case I, the tracking error converges asymptotically to zero with state dependent threshold, i.e., the worst case sampled control policy, for the sampling condition in (23). However, enforcing a minimum constant threshold, r, for the sampling condition, as shown in Case II of the proof, one can only show ultimate boundedness of the tracking error and the bound is a function of r. Therefore, the minimum threshold r can be selected arbitrarily close to zero. Heuristically, it is seen from simulation that the state dependent threshold in (23) does not go to zero before the convergence of system state.

From the optimal value perspective, the next corollary quantifies the degree of the optimality by computing the optimal value for the event-triggered implementation when compared to the continuous saddle point optimal value.

Corollary 1. *Let the hypothesis of Theorem 1 hold. Then, the optimal cost for the event-triggered implementation with policy w_s^* is given by*

$$J(;, w_s^*) = V^*(\chi) + \int_t^\infty [e_u^T(R - \gamma^2 I)e_u - \gamma^2 e_u^{*T}e_u^*]d\tau. \tag{24}$$

Proof. Refer to Appendix.

Remark 5. The novel performance index in (14) provides an extra degree of freedom, in terms of the parameter γ, in optimizing the value when compared to the optimal value in Vamvoudakis et al. (2017). In addition, when compared to the threshold coefficient parameter σ in the traditional event-triggering condition in Tabuada (2007), this penalizing term γ as threshold coefficient in the sampling condition (23) also determines the degree of optimality.

3.3. Zeno-free behavior of system

To show the sampling instants are not accumulated, i.e., Zeno-free behavior of the system, we will use a more conservative event-sampling condition. Define the augmented state sampling error e_s as the error between the continuous augmented state $\chi(t)$ and the sampled augmented state $\chi_s(t)$. It can be expressed as

$$e_s(t) = \chi_s(t) - \chi(t), \quad \forall t \in [t_k, t_{k+1}). \tag{25}$$

The following standard assumption and the technical lemma are necessary to proceed.

Assumption 5. The optimal policies u^* and e_u^* are locally Lipschitz in a compact set, such that, $\|u_s^* - u^*\| = \|e_u\| \leq L_u \|e_s\|$ and $\|e_u^*(\chi) - e_u^*(\chi_s)\| \leq L_u \|e_s\|$ where $L_u > 0$ is the Lipschitz constant.

Lemma 2. *If the inequality $L_u\|e_s(t)\| \leq (1/4\gamma^2)\|G^T(\chi_s)V_{\chi_s}^*\|$ holds, then the inequality $e_u^T(t)e_u(t) \leq (1/4\gamma^4)V_\chi^{*T}G(\chi)G^T(\chi)V_\chi^*$ also holds.*

Proof. By definition we have $\frac{1}{4\gamma^2}\|G^T(\chi_s)V^*_{\chi_s}\| = \frac{1}{2}\|e^*_u(t_k)\|$ and, therefore, the expression $L_u\|e_s(t)\| \le (1/4\gamma^2)\|G^T(\chi_s)V^*_{\chi_s}\|$ can be rewritten as $2L_u\|e_s(t)\| \le \|e^*_u(t_k)\| = \|e^*_u(t) + e^*_u(t_k) - e^*_u(t)\| \le \|e^*_u(t)\| + \|e^*_u(t_k) - e^*_u(t)\| \le \|e^*_u(t)\| + L_u\|e_s(t)\|$.

Rearranging the expression, leads to

$$L_u\|e_s(t)\| \le \|e^*_u(t)\|. \tag{26}$$

By Assumption 5, $\|e_u(t)\| < L_u\|e_s(t)\|$, comparing with (26), it holds that $\|e_u(t)\| \le \|e^*_u(t)\|$. Squaring both sides it holds that $e^T_u(t)e_u(t) \le (1/4\gamma^4)V^{*T}_\chi G(\chi)G^T(\chi)V^*_\chi$.

Corollary 2. *Let the hypothesis of Theorem 1 hold. Then, the sampling condition defined by*

$$t_{k+1} = \inf\{t > t_k \mid L_u\|e_s(t)\| = \max\{r, \frac{1}{4\gamma^2}\|G^T(\chi_s)V^*_{\chi_s}\|\}\}, \tag{27}$$

ensures the tracking error converges to it's ultimate bound. Further, the minimum inter-sample time $\tau_m = \inf_{k\in\{0,\mathbb{N}\}}(\tau_k) = \inf_{k\in\{0,\mathbb{N}\}}(t_{k+1} - t_k) > 0$ *where* τ_k *is given by*

$$\tau_k > \frac{1}{\kappa_1}\ln\left(\frac{\kappa_1 r}{\kappa_2} + 1\right). \tag{28}$$

where $\kappa_1 = G_M L_u + L_u\varrho$ and $\kappa_3 = L_u\varrho\|\chi_s\|$.

Proof. Refer to Appendix.

Since, an analytical closed form solution to the HJI equation (22) is difficult to compute (Bellman, 2013), the solution is approximated using a NN and is presented in the next section.

4. Approximate solution for the min–max optimization problem

In this section, the optimal value function, which is the solution of the HJI equation is approximated using a functional link NN to design the optimal control policy and the sampling condition.

4.1. Value function approximation and event-sampled bellman error

Recalling Assumption 4, the solution to the HJI equation, i.e., the optimal value function $V^*(\chi)$ is smooth and continuous. By the universal approximation property (Lewis et al., 1998), the value function can be approximated using neural network, referred to as *critic NN*, in a compact set $\Omega_\chi \subset \Re^{2n}$. Alternatively, in a compact set Ω_χ, there exist an ideal weight vector W and a basis function $\phi(\chi)$ such that the value function can be expressed as

$$V^*(\chi) = W^T\phi(\chi) + \varepsilon(\chi) \tag{29}$$

where $W \in \Re^{l_o}$ is the target weight vector, which is unknown, $\phi : \Re^{2n} \to \Re^{l_o}$ is the activation function, and $\varepsilon : \Re^{2n} \to \Re$ is the approximation error. Further, the activation function satisfies $\phi(0) = 0$ and l_o is the number of neurons in the hidden layer. The following standard assumption for the NN is used for analysis.

Assumption 6. The unknown target weight vector W, the activation function ϕ, and the reconstruction error ε are bounded in the compact sets. This means, there exist constants $W_M > 0$, $\phi_M > 0$ and $\varepsilon_M > 0$ such that, $\|W\| \le W_M$, $\sup_{\chi\in\Omega_\chi}\|\phi(\chi)\| \le \phi_M$, and $\sup_{\chi\in\Omega_\chi}\|\varepsilon(\chi)\| \le \varepsilon_M$. Further, the gradient of approximation error satisfies $\sup_{\chi\in\Omega_\chi}\|\nabla\varepsilon(\chi)\| \le \bar{\varepsilon}_M$ where $\bar{\varepsilon}_M > 0$ is a constant and $\nabla\varepsilon(\chi) = \frac{\partial\varepsilon(\chi)}{\partial\chi}$.

The optimal mismatch control policy in a parametric form using NN approximation can be computed as

$$w^*(\chi) = -(1/2)R^{-1}G^T(\chi)(\nabla\phi^T(\chi)W + \nabla\varepsilon(\chi)), \tag{30}$$

and, similarly, the worst case threshold becomes

$$e^*_u(\chi) = -(1/2\gamma^2)G^T(\chi)(\nabla\phi^T(\chi)W + \nabla\varepsilon(\chi)), \tag{31}$$

where $\nabla\phi(\chi) = \frac{\partial\phi(\chi)}{\partial\chi}$.

The estimated value function $\hat{V}(\chi)$ is given by

$$\hat{V}(\chi) = \hat{W}^T\phi(\chi_s), \quad \forall t \in [t_k, t_{k+1}), \tag{32}$$

where $\hat{W} \in \Re^{l_o}$ is NN weight estimates and $\phi(\chi_s) \in \Re^{l_o}$ is the activation function with sampled augmented state χ_s as input. Then, estimated mismatch control policy

$$w(\chi) = -(1/2)R^{-1}G^T(\chi)(\nabla\phi^T(\chi_s)\hat{W}) \tag{33}$$

and the estimated control input can be expressed as

$$u = -\frac{1}{2}R^{-1}G^T(\chi)\nabla\phi^T(\chi_s)\hat{W} + g^+(x_d)(\zeta(x_d) - f(x_d)). \tag{34}$$

The estimated sampled control policy now can be expressed as

$$\begin{aligned} u_s(t) &= -(1/2)R^{-1}G^T(\chi_s)\nabla\phi^T(\chi_s)\hat{W} \\ &\quad + g^+(x_{ds})(\zeta(x_{ds}) - f(x_{ds})), \forall t \in [t_k, t_{k+1}). \end{aligned} \tag{35}$$

Further, the estimated threshold is given by

$$\hat{e}_u(t) = (1/2\gamma^2)G^T(\chi)\nabla\phi^T(\chi_s)\hat{W}. \tag{36}$$

To update the weights of the critic NN we will use the Bellman principle of optimality. Note that, the value function (14) in the ETC frame work can equivalently be expressed as

$$J(\chi(0)) = \sum_{k=0}^{\infty}\left(\int_{t_k}^{t_k+\tau_k}(\chi^T\bar{Q}\chi + w^TRw - \gamma^2\hat{e}^T_u\hat{e}_u)d\tau\right).$$

By Bellman principle of optimality (Bellman, 2013), the Bellman equation in an integral from is given by

$$V^*(t_{k+1}) - V^*(t_k) = \int_{t_k}^{t_{k+1}}\left(-\chi^T\bar{Q}\chi - w^TRw + \gamma^2\hat{e}^T_u\hat{e}_u\right)d\tau \tag{37}$$

where $t_{k+1} = t_k + \tau_k$. The event-based integral Bellman equation (37), with the critic NN (29), yields

$$W^T\Delta\phi(\tau_k) + \Delta\varepsilon(\tau_k) = \int_{t_k}^{t_{k+1}}(-\chi^T\bar{Q}\chi - w^TRw + \gamma^2\hat{e}^T_u\hat{e}_u)d\tau \tag{38}$$

where $\Delta\phi(\tau_k) = \phi(\chi(t_{k+1})) - \phi(\chi(t_k))$ and $\Delta\varepsilon(\tau_k) = \varepsilon(\chi(t_{k+1})) - \varepsilon(\chi(t_k))$.

With estimated critic NN weights (32), the Bellman error can be expressed as

$$\begin{aligned} \delta_{k+1} &= \int_{t_k}^{t_{k+1}}(\chi^T\bar{Q}\chi + w^TRw - \gamma^2\hat{e}^T_u\hat{e}_u)d\tau + \hat{V}(t_{k+1}) - \hat{V}(t_k) \\ &= \int_{t_k}^{t_{k+1}}(\chi^T\bar{Q}\chi + w^TRw - \gamma^2\hat{e}^T_u\hat{e}_u)d\tau + \hat{W}^T\Delta\phi(\tau_k) \end{aligned} \tag{39}$$

where δ_{k+1} is the Bellman residual error or temporal difference error calculated at the occurrence of $k+1$ event.

The critic NN weights are learned online such that the Bellman residual error (39) is minimized. Since the augmented system states at the controller are updated only at the sampling instants, the Bellman residual error (39) can only be computed at the sampling instants. Therefore, the NN weights are updated as a jump in the weight at the triggering instants $t_k, \forall k \in \{0, \mathbb{N}\}$ with the new state feedback information, given as

$$\hat{W}^+ = \hat{W} - \alpha_2 \frac{\Delta\phi(\tau_{k-1})}{(1+\Delta\phi^T(\tau_{k-1})\Delta\phi(\tau_{k-1}))^2}\delta_k^T, \quad t = t_k \tag{40}$$

where $\hat{W}^+ = \hat{W}(t_k^+)$ where t_k^+ is the time instant just after t_k. Further, to utilize the inter-sample times and accelerate the convergence of NN weights, the Bellman residual error computed at the previous sampling instants t_k is used to update the critic NN weight estimates. The continuous update (flow) during the inter-sample times is defined as

$$\dot{\hat{W}} = -\alpha_1 \frac{\Delta\phi(\tau_{k-1})}{(1+\Delta\phi^T(\tau_{k-1})\Delta\phi(\tau_{k-1}))^2}\delta_k^T, \quad t \in (t_k, t_{k+1}) \tag{41}$$

where $\delta_k = \int_{t_{k-1}}^{t_k}(\chi^T\bar{Q}\chi + w^TRw - \gamma^2\hat{e}_u^T\hat{e}_u)d\tau + \hat{W}^T\Delta\phi(\tau_{k-1})$ is the Bellman residual error at t_k, derived from (39) where $\alpha_1 > 0$, and $\alpha_2 > 0$ are learning gains.

Remark 6. Note that the integration $\int_{t_{k-1}}^{t_k}(\cdot)d\tau$ and the difference $\Delta\phi(\tau_k)$ can be computed using the augmented state χ and mismatch control input w information at two consecutive sampling instants t_k and t_{k-1}, which are available at the controller at the kth-time instant.

Remark 7. The update laws in (40) and (41) can be referred to as impulsive parameter update scheme with (40) as jump and (41) as flow dynamics. Further, the update during the inter-sample times is motivated by the traditional value/policy iteration based ADP schemes (Bertsekas et al., 1995). This ensures boundedness of the closed-loop system parameters during the flow periods (refer to the proof of Theorem 2) when the control policy is not updated.

4.2. Adaptive event-sampling condition and stability

The event-based sampling condition with estimated $\hat{e}_u$ in (36) can be defined as

$$\begin{aligned} t_{k+1} = \inf\{t > t_k \mid & e_u(t)^T e_u(t) \\ &= \max(r^2, \frac{1}{4\gamma^4}\hat{W}^T\nabla\phi(\chi_s)G(\chi)G^T(\chi)\nabla\phi^T(\chi_s)\hat{W})\}. \end{aligned} \tag{42}$$

Defining the NN weight estimation error $\tilde{W} = W - \hat{W}$, the Bellman residual error, by subtracting (39) from (38) with an event-step backward, can be represented as

$$\delta_k = -\tilde{W}^T\Delta\phi(\tau_{k-1}) - \Delta\varepsilon(\tau_{k-1}). \tag{43}$$

The Bellman residual error δ_k using (43) is not computable since the target NN weight W is unknown and will only be used for demonstrating the stability; presented in the next theorem.

The event-sampled augmented tracking error system, by defining a concatenated state vector $\xi = [\chi^T, \tilde{W}^T]^T \in \Re^{2n+l_o}$, can be expressed as a nonlinear impulsive dynamical system as

$$\dot{\xi} = \begin{bmatrix} F(\chi) + G(\chi)w + G(\chi)e_u \\ \alpha_1 \dfrac{\Delta\phi(\tau_{k-1})}{(1+\Delta\phi^T(\tau_{k-1})\Delta\phi(\tau_{k-1}))^2}\delta_k^T \end{bmatrix}, \ \xi \in \mathcal{C}, \quad t \in (t_k, t_{k+1}) \tag{44}$$

and

$$\xi^+ = \begin{bmatrix} \chi \\ \tilde{W} + \alpha_2 \dfrac{\Delta\phi(\tau_{k-1})}{(1+\Delta\phi^T(\tau_{k-1})\Delta\phi(\tau_{k-1}))^2}\delta_k^T \end{bmatrix}, \ \xi \in \mathcal{D}, \quad t = t_k. \tag{45}$$

where (44) are the dynamics of the system during the inter-sample times, referred to as flow dynamics, and (45) are the dynamics at the sampling instants referred to as jump dynamics. The sets

$$\begin{aligned}\mathcal{C} \triangleq \{\xi \in \Re^{2n+l_o} \mid & e_u(t)^T e_u(t) \\ &< \max(r^2, \frac{1}{4\gamma^4}\hat{W}^T\nabla\phi(\chi_s)G(\chi)G^T(\chi)\nabla\phi^T(\chi_s)\hat{W})\}\end{aligned}$$

and

$$\begin{aligned}\mathcal{D} \triangleq \{\xi \in \Re^{2n+l_o} \mid & e_u(t)^T e_u(t) \\ &\geq \max(r^2, \frac{1}{4\gamma^4}\hat{W}^T\nabla\phi(\chi_s)G(\chi)G^T(\chi)\nabla\phi^T(\chi_s)\hat{W})\}\end{aligned}$$

are the flow and jump sets, respectively.

For brevity and to facilitate the proof of the theorem, presented next, the following variables are defined.

$$\mathrm{g}_r = GR^{-1}G^T, \quad \mathrm{g}_\phi = \nabla\phi GR^{-1}G^T\nabla^T\phi, \quad \mathrm{g}_\gamma = \nabla\phi GG^T\nabla^T\phi. \tag{46}$$

Further, using Assumptions 2 and 6, the following bounds are developed in compact set Ω_χ:

$$\begin{aligned}
&\|\frac{1}{2}W^T\mathrm{g}_\phi + \frac{1}{2}\nabla\varepsilon^T\mathrm{g}_r\nabla\phi^T\| = \iota_1, \quad \|\mathrm{g}_\gamma\| = \iota_2, \\
&\frac{1}{2}\|W^T\nabla\phi\mathrm{g}_r\nabla\varepsilon\| + \frac{1}{2}\|\nabla\varepsilon^T\mathrm{g}_r\nabla\varepsilon\| = \iota_3, \|\mathrm{g}_r\| = \iota_4, \\
&\|\mathrm{g}_\phi\| = \iota_5, \quad (\iota_1 + \frac{\iota_2}{2\gamma^2}) = \varpi_1, \frac{\iota_2}{2\gamma^2}\|W\|^2 + \iota_3 = \varpi_2, \\
&\varpi_3 = \varpi_2 + \frac{\alpha}{4}\Delta\varepsilon_M^2, \quad \varpi_5 = \frac{1}{2}\alpha^2\Delta\varepsilon_M^2 + \frac{1}{2}\alpha\Delta\varepsilon_M^2 + \frac{1}{2}\alpha^2\Delta\phi_M\Delta\varepsilon_M^2.
\end{aligned} \tag{47}$$

The facts

$$\|\frac{\Delta\phi(\tau_{k-1})}{1+\Delta\phi^T(\tau_{k-1})\Delta\phi(\tau_{k-1})}\| \leq \frac{1}{2}, \ \frac{1}{1+\Delta\phi^T(\tau_{k-1})\Delta\phi(\tau_{k-1})} \leq 1 \tag{48}$$

for every vector $\Delta\phi$ are also utilized to claim the practical stability using Lyapunov analysis.

Before presenting the main results, the following standard assumption for persistence of excitation (PE) condition for parameter estimation (Ioannou & Fidan, 2006) is presented for completeness.

Assumption 7. The vector $\varphi(\tau) \triangleq \frac{\Delta\phi(\tau)}{1+\Delta\phi^T(\tau)\Delta\phi(\tau)} \in \Re^{l_o}$ is persistently exciting, i.e., there exist a time period $T > 0$ and a constant $\varphi_\delta > 0$ such that over the interval $[t, t+T]$ the regressor vector satisfies $\int_t^{t+T}\varphi(\tau)\varphi^T(\tau)d\tau \geq \varphi_\delta I$, where I is the identity matrix of appropriate dimensions.

Theorem 2. *Consider the event-triggered system represented as a nonlinear impulsive hybrid dynamical system in (44) and (45) with control policy (35). Suppose Assumptions 1–7 hold, the NN initial weights $\hat{W}(0)$ initialized in a compact set Ω_W and the initial control policy be admissible. Then, there exists a positive integer $N > 0$, such that, the tracking error e_r and the NN weight estimation error $\tilde{W}$, are locally ultimately bounded for all sampling instants $t_k > t_N$, provided event-based sampling instants are obtained using (42) and the weight tuning gains are selected as $0 < \alpha_1 < 1$ and $0 < \alpha_2 < \frac{1}{2}$. Further, $\|V^* - \hat{V}\|$ and $\|w_s - w_s^*\|$ are also ultimately bounded.*

Proof. Refer to Appendix.

Corollary 3. *Let the hypothesis of Theorem 2 hold. Then, the set of the sequence of event sampling instants $\{t_k \mid k \in \{0, \mathbb{N}\}, e_u^T e_u$*

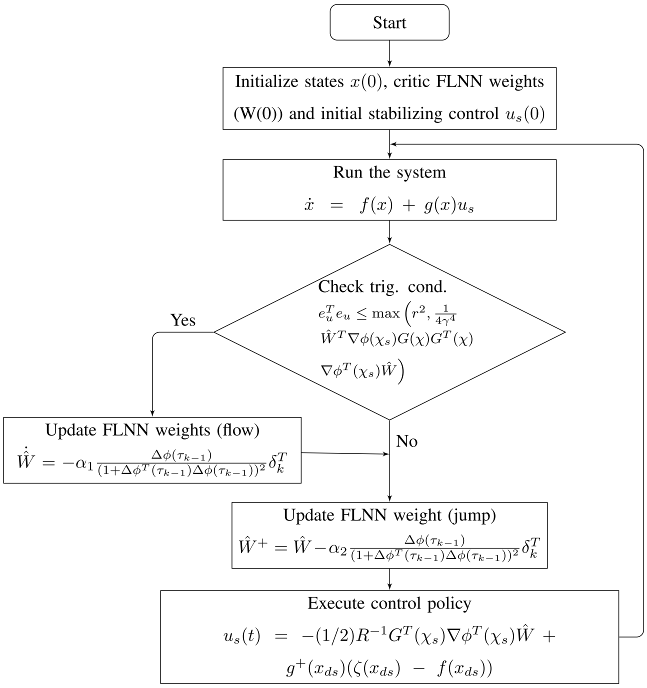

Fig. 1. Implementation of the proposed algorithm.

$= \max(r^2, \frac{1}{4\gamma^4}\hat{W}^T\nabla\phi(\chi_s)GG^T(\chi)\nabla^T\phi(\chi_s)\hat{W})\}$ *determined by* (42) *converges the closed neighborhood of the actual sampling sequence* $\{t_k \mid k \in \{0, \mathbb{N}\}, e_u(t)^T e_u(t) = \max\{r^2, (1/4\gamma^4)V_\chi^{*T} GG^T V_\chi^*\}\}$ *determined in* (23).

Proof. The proof is the direct result of Theorem 2 due to the convergence of $\|\hat{V} - V^*\|$ to its ultimate bound.

Corollary 4. *Let the hypothesis of Theorem 2 hold. Then, with the event-based sampling condition*

$$t_{k+1} = \inf\{t > t_k \mid L_u\|e_s(t)\| = \max\{r, \frac{1}{4\gamma^2}\|G^T(\chi_s)\hat{V}_{\chi_s}\|\}\} \qquad (49)$$

the tracking error e_r *and NN weight estimation error* $\tilde{W}$ *are locally ultimately bounded. Further, the minimum inter-sample time* $\tau_m = \inf_{k\in\{0,\mathbb{N}\}}(\tau_k) > 0$ *where* τ_k *is given by*

$$\tau_k > \frac{1}{\kappa_1}\ln\left(\frac{\kappa_1 r}{\kappa_2} + 1\right). \qquad (50)$$

Proof. The proof follows the same lines of arguments as in Lemma 2 and Corollary 2 and, therefore, omitted.

Next, the numerical simulation results are presented.

5. Implementation and simulation results

The implementation of the proposed analytical design is detailed in the flow chart shown in Fig. 1.

A numerical example of Van-der-Pol oscillator, whose dynamics is given by

$$\dot{x} = f(x) + g(x)u,$$

with $f = \left[x_2, \quad x_2(x_1^2 - 1) + x_1\right]^T$ and $g = [0, \quad 1]^T$, is considered for the validation using simulation.

The reference trajectory considered is given by $x_d = [0.5cos(2t), \quad -sin(2t)]^T$ with $\zeta(x_d) = [x_{d2}, \quad -4x_{d1}]^T$. The penalty matrices in the performance index (14) are selected as $Q = 50 \times diag[1, 0.1]$, $R = 0.1$, $\gamma = 0.35$. The NN weights are initialized randomly from a uniform distribution in the interval $[0, \ 2]$. The learning gains are selected as $\alpha_1 = 0.045$, and $\alpha_2 = 0.08$. A polynomial regression vector $\phi(\chi) = (1/8)[\chi_1^2, \chi_1\chi_2, \chi_1\chi_3, \chi_1\chi_4, \chi_2^2, \chi_2\chi_3, \chi_2\chi_4, \chi_3^2, \chi_3\chi_4, \chi_1^2\chi_2, \chi_1^2\chi_3, \chi_1^2\chi_4, \chi_2^2, \chi_2^2\chi_3, \chi_2^2\chi_4, \chi_3^2\chi_4, \chi_1^2\chi_2^2, \chi_1^2\chi_3^2, \chi_1^2\chi_4^2, \chi_2^2\chi_3^2, \chi_2^2\chi_4^2, \chi_3^2\chi_4^2, \chi_4^2]^T$ is used for approximating the solution of the HJI equation. A normally distributed probing noise in the interval $[0, 1]$ is added to the regressor vector for satisfying PE condition to ensure the convergence of the NN weights. The simulation is run for 35 s with initial states $x_0 = [2, \quad 1.5]^T$.

The results of the proposed near optimal event-based ADP based scheme are compared with sampled data implementation of the scheme, where the feedback signals are sampled and the controller is executed periodically. The plots are super imposed to verify the effectiveness of the proposed approach. The notation (e) and (c) next to the simulation variables in the figures indicate event-based implementation and sampled data implementation, respectively. All simulation parameters including the initial conditions are kept same for both the cases. Learning gain for the NN weight update during the flow period, which is not available in the sampled data case, is adjusted to obtain a comparable state convergence rate for fair comparison in terms of computation and performance cost.

The event-sampling condition in (40) with $r = 0.005$ is utilized for the sampling and controller execution. The states and the desired trajectory are shown in Fig. 2. It clear from the figure that with the event based implementation of the control input closely follows the sampled data implementation with less frequent control execution. The system states converge arbitrarily close to the desired states in about 20 s. The large initial tracking error is due to the approximated controller with randomly initialized NN weights. As the NN weight estimates converge close to the target values, the system states converge close to the desired trajectories. The convergence of the tracking errors for both event-based and sampled data implementation to their respective bounds are shown in Fig. 3. Note that the event-based tracking error closely follows the sampled-data tracking error. This is achieved by selecting the NN weight tuning gain during the flow period appropriately, which accelerates the convergence rate of the weight updates.

The convergence of the event-based Bellman residual error close to zero is shown in Fig. 4. Note that the Bellman error converges earlier than the state convergence in both the cases. This implies the approximate value function converges to arbitrary close neighborhood of their optimal values before the states converge to their respective desired states. It is observed that the Bellman error for the event-based implementation case has spikes before convergence when compared to the smooth plot of sampled data implementation. This is due to the aperiodic control execution. However, the Bellman error converges close to zero as in the case of sampled data implementation. The cost comparison is shown in Fig. 5. The cumulative cost for event-based approach is close to each other. Note that the penalty factor γ in (12) determines the degree of optimality (Corollary 1) and different values of γ will lead to different costs.

From the reduction of computation point of view, the cumulative number of sampling instants and the inter-sampling times are shown in Fig. 6(a) and (b). It is observed that the control is executed 51.18% of times when compared to the periodic sampled data execution for $\gamma = 0.35$ during the simulation time of 35 s. This shows a 48.82% reduction of feedback communication and computational load. Note that γ is selected such that $\gamma^2 > \lambda_{max}(R)$ where $\lambda_{max} = 0.1$. Different values of γ and r will result in different number of sampling. From Fig. 6(b) it is evident that the

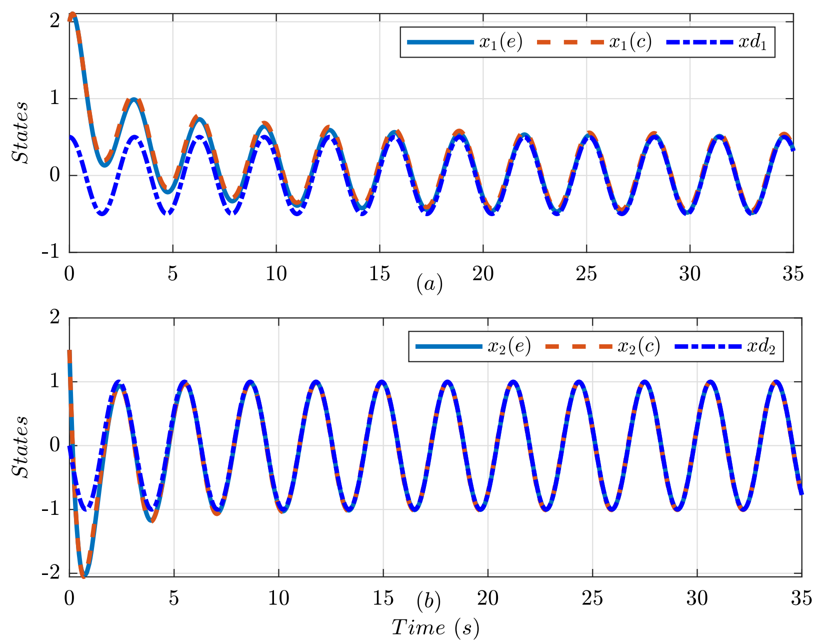

Fig. 2. System state and desired trajectories.

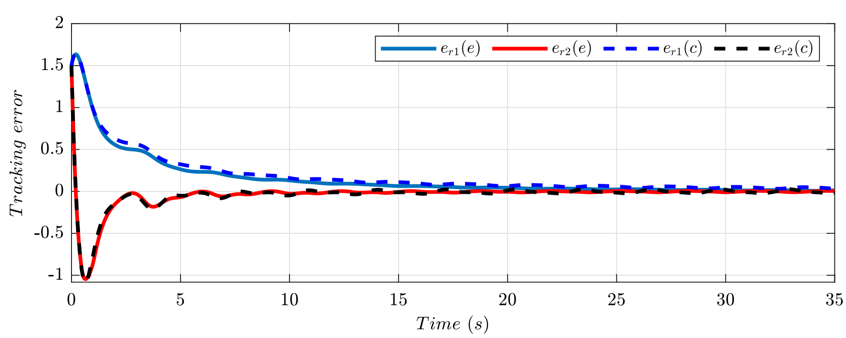

Fig. 3. Convergence of tracking errors.

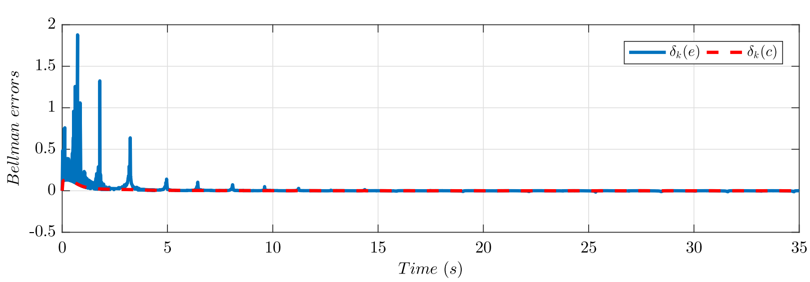

Fig. 4. Convergence of event-based Bellman errors.

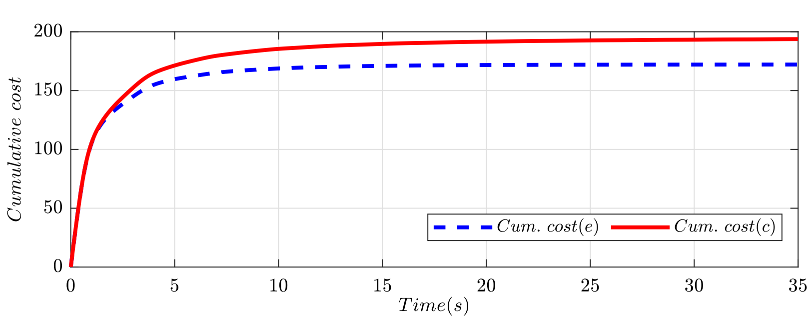

Fig. 5. Comparison of cumulative costs.

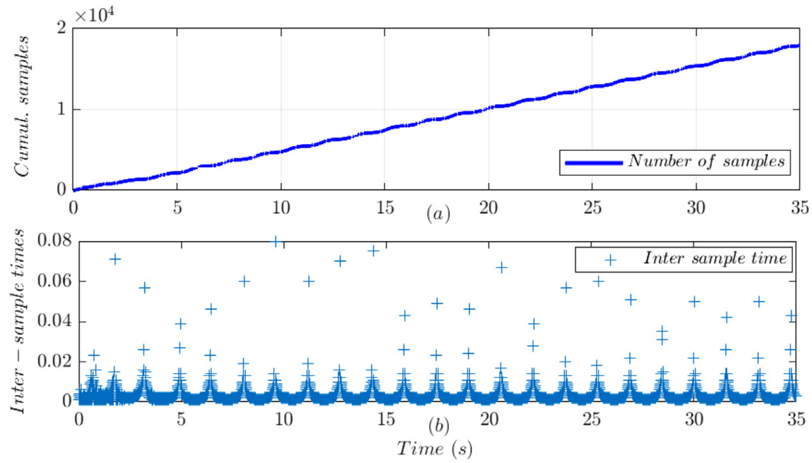

Fig. 6. Cumulative number of sampling instants and inter-sample times.

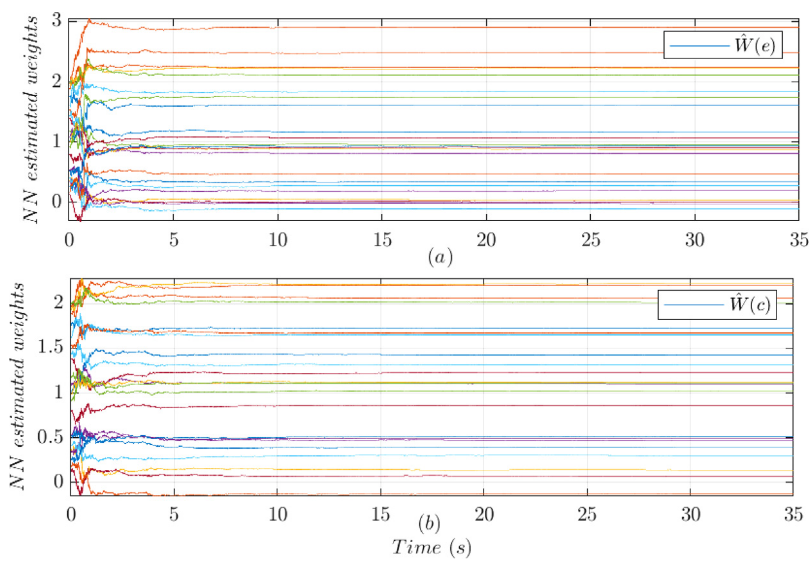

Fig. 7. Convergence of estimated NN weights.

sampling instants are aperiodic and the events are crowded till 2.5 s. This is due to the high initial tracking error in the learning phase of the NN. Once the NN estimated weights converge close to the target values, shown in Fig. 7, the sampling instants are reduced.

Since the ideal NN weights are unknown, it is not possible to show convergence of the NN estimated weight to the target values. However, from Fig. 7(a) and (b), it can be observed that the NN weight estimates reach its steady state, i.e., become constant, approximately about 18 s for both the cases. Therefore, it can be concluded that the NN weights converge to a close neighborhood of the target weights. Alternatively, the NN weight estimation error is ultimately bounded. Note that the target values for both the cases are different since the optimal value function for sampled data case is different from the proposed event-based case.

6. Conclusions

A near optimal adaptive sampling scheme is proposed for state trajectory tracking control of nonlinear systems. An approximate saddle point solution of the HJI equation is obtained using NN based approximation and impulsive learning of NN weights. The impulsive weight update scheme maintained boundedness of the closed-loop system parameters during inter-sample times and also accelerated the convergence rate. The triggering condition is designed using the worst case threshold for sampled error policy, which maximized the sampling intervals and reduced the controller execution by 48.82%. Moreover, the sampling intervals are near optimal along with the control policy. The analytical design is also verified using numerical example and the results are also compared with its sampled data counter part. It is observed that for a similar state trajectory, the saddle point solution of the event-based implementation saved the computation and cost. As presented, the optimal event-based control can also save communication resources. Our future work will include the event-based sampling and control for nonlinear networked control systems with network imperfections.

Appendix

Proof of Lemma 1. Consider a continuously differentiable positive definite candidate Lyapunov function $L_\chi : \Re^n \times \Re_{\geq 0} \to \Re$ given by $L_\chi(e_r, t) = V_t^*(e_r, t)$. The time derivative along the system dynamics (13) can be obtained as

$$\dot{L}_\chi = V_\chi^{*^T} \dot{\chi} = V_\chi^{*^T}(F + Gw + Ge_u).$$

Replacing w with its optimal policy w^* and adding and subtracting $V_\chi^{*T} G e_u^*$, the first derivative

$$\dot{L}_\chi = V_\chi^{*T}(F + Gw^* + Ge_u^*) + V_\chi^{*T} G(e_u - e_u^*). \tag{51}$$

From the HJI equation in (22), we have $-\chi^T \bar{Q}\chi - w^{*T} R w^* + \gamma^2 e_u^{*T} e_u^* = V_\chi^{*T}[F + Gw^* + Ge_u^*]$. Substituting in (51), the first difference results in

$$\dot{L}_\chi = -\chi^T \bar{Q}\chi - w^{*T} R w^* + \gamma^2 e_u^{*T} e_u^* + V_\chi^{*T} G(e_u - e_u^*). \tag{52}$$

From (18), we have $G^T V_\chi^* = 2\gamma^2 e_u^*$. Inserting the relation in (52) leads to

$$\dot{L}_\chi = -(\chi^T \bar{Q}\chi + w^{*T} R w^* + \gamma^2 e_u^{*T} e_u^*) + 2\gamma^2 e_u^{*T} e_u.$$

Using the definitions of the augmented state χ and $\bar{Q}$ and Young's inequality $2a^T b \le a^T a + b^T b$, the first difference can be expanded as

$$\dot{L}_\chi \le -(e_r^T Q e_r + w^{*T} R w^*) + \gamma^2 e_u^T e_u \le -\delta_1(\|e_r\|) + \delta_2(\|e_u\|) \tag{53}$$

where $\delta_1(\|e_r\|) = e_r^T Q e_r > 0$ since Q is symmetric positive definite and $\delta_2(\|e_u\|) = \gamma^2 e_u^T e_u$. From (53), $\dot{L}_\chi < 0$ out side the set $B_{e1} \triangleq \{e_r \in \Re^n | \|e_r\| > \delta_1^{-1}(\delta_2(\|e_u\|))\}$. Thus, it can be concluded that L_χ is a local ISS Lyapunov function (Khalil, 1996) and the augmented system (13) is ISS with respect to e_u.

Proof of Theorem 1. Consider a continuously differentiable positive definite candidate Lyapunov function $L_\chi : \Re^n \times \Re_{\ge 0} \to \Re$ as in Lemma 1. The time derivative along the system dynamics (13) can be obtained from (53) as

$$\dot{L}_\chi \le -(e_r^T Q e_r + w^{*T} R w^*) + \gamma^2 e_u^T e_u \tag{54}$$

Based on the sampling condition (23), the following two cases arise.

Case I: For the triggering condition $e_u^T e_u \le (1/4\gamma^4) V_\chi^{*T} G G^T V_\chi^*$
The Lyapunov's first derivative is upper bounded by

$$\dot{L}_\chi \le -(e_r^T Q e_r + w^{*T} R w^*) + \frac{1}{4\gamma^2} V_\chi^{*T} G G^T V_\chi^*. \tag{55}$$

From (17), we have $-2Rw^* = G^T V_\chi^*$. Using the relation the first derivative

$$\dot{L}_z \le -(e_r^T Q e_r + w^{*T}[R - \frac{1}{\gamma^2} R^2] w^*) \le -\delta_3(\|e_r\|) \tag{56}$$

where $\delta_3(\|e_r\|) = e_r^T Q e_r + w^{*T}[R - \frac{1}{\gamma^2} R^2] w^* > 0$. By selecting $\gamma^2 > \lambda_{max}(R)$, from (56), the Lyapunov's first derivative is negative. By Lyapunov theorem, the tracking error converges asymptotically to zero.

Case II: For the triggering condition $e_u^T e_u \le r^2$
The Lyapunov's first derivative is upper bounded by

$$\dot{L}_\chi \le -\lambda_{min}(Q)\|e_r\|^2 - w^{*T} R w^* + \gamma^2 r^2. \tag{57}$$

Defining $\lambda_{qr} = min\{\lambda_{min}(Q), \lambda_{min}(R)\}$, from (57), the Lyapunov first derivative $\dot{L}_\chi$ is negative out side the set $B_{e2} = \{e_r \in \Re^n | \|e_r\|^2 > \frac{\gamma^2 r^2}{\lambda_{qr}}\}$. Consequently the tracking error e_r is ultimately bounded.

Finally, by combining case I and case II, the tracking error e_r is ultimately bounded with a bound B_{e2}.

Proof of Corollary 1. Consider the cost-function (14). With the event sampled mismatch control input w_s^*, the performance cost

$$J(;, w_s^*) = V^*(\chi) + \int_t^\infty [H_e(\chi, w_s^*, V_\chi^*) - H^*] d\tau \tag{58}$$

where $H_e(\cdot)$ is the event-based Hamiltonian. Recall that the Hamiltonian function $H^*(\chi, w^*, e_u^*) = 0$. Using the Hamiltonian (22) and the control policies (17) one has

$$\begin{aligned} H^* - H_e &= w^{*T} R w^* - \gamma^2 e_u^{*T} e_u^* + V_\chi^{*T}[Gw^* + Ge_u^*] \\ &- w_s^{*T} R w_s^* + \gamma^2 e_u^T e_u - V_\chi^{*T} G w_s^* \end{aligned} \tag{59}$$

where $e_u = u_s^* - u^*$. From the definition of the control policies, we have $G^T V_\chi^* = -2Rw^*$. Using this in (59), $H^* - H_e = -[w^{*T} R w^* + w_s^{*T} R w_s^* - 2w^{*T} R w_s^*] + \gamma^2[e_u^{*T} e_u^* + e_u^T e_u]$. Completing the squares, we have

$$H^* - H_e = (w^* - w_s^*)^T R(w - w_s^*) + \gamma^2[e_u^{*T} e_u^* + e_u^T e_u] \tag{60}$$

Substituting $H^* = 0$, reveals

$$H_e = e_u^T R e_u - \gamma^2[e_u^{*T} e_u^* + e_u^T e_u] \tag{61}$$

Substituting (61) in (58), reveals (24).

Proof of Corollary 2. To show the ultimate boundedness of the tracking error by invoking Theorem 1, we will show that if the event-based sampling condition (27) is satisfied, then the sampling condition (23) is also satisfied. The proof is presented in two cases resulting from the sampling condition (27).

Case I: For the case $r < \frac{1}{4\gamma^2}\|G^T(\chi_s) V_{\chi_s}^*\|$
The sampling condition in (27), for stability, can be expressed $L_u\|e_s\| \le \frac{1}{4\gamma^2}\|G^T(\chi_s) V_{\chi_s}^*\|$. From Lemma 1, this implies $e_u^T e_u \le (1/4\gamma^4) V_\chi^{*T} G(\chi) G^T(\chi) V_\chi^*$.

Further, from the actual triggering condition (23), if $r^2 > (1/4\gamma^4) V_\chi^{*T} G(\chi) G^T(\chi) V_\chi^*$, it holds that $e_u^T e_u \le (1/4\gamma^4) V_\chi^{*T} G(\chi) G^T(\chi) V_\chi^* < r^2$.

Case II: For the case $r > \frac{1}{4\gamma^2}\|G^T(\chi_s) V_{\chi_s}^*\|$
The sampling condition (27), for stability, can be expressed $L_u\|e_s\| \le r$. From Assumption 4, $\|e_u(t)\| < L_u\|e_s(t)\|$, and squaring both sides we have $e_u^T e_u \le r^2$.

Further, from the actual triggering condition (23), if $r^2 > (1/4\gamma^4) V_\chi^{*T} G(\chi) G^T(\chi) V_\chi^*$, the triggering condition is trivially satisfied. For the case $r^2 < (1/4\gamma^4) V_\chi^{*T} G(\chi) G^T(\chi) V_\chi^*$, it holds that $e_u^T e_u \le r^2 \le (1/4\gamma^4) V_\chi^{*T} G(\chi) G^T(\chi) V_\chi^*$.

Consequently, from both cases, the conservative triggering condition (27) implies actual triggering condition (23). Therefore, by Theorem 1 the tracking error is ultimately bounded.

To show the positive inter-sample times using the triggering condition (27), define $\bar{e}_s \triangleq L_u\|e_s\|$ and taking the time-derivative reveals

$$\frac{d}{dt}\bar{e}_s \le \|\dot{\bar{e}}_s\| = L_u\|\dot{\chi}_s - \dot{\chi}\|, \quad t \in [t_k, t_{k+1}). \tag{62}$$

Along the system dynamics (14), the derivative can be expressed as

$$\left\|\dot{\bar{e}}_s\right\| = L_u\|\dot{\chi}\| = L_u\left\|F(\chi) + G(\chi)w^* + G(\chi)e_u\right\|, \quad t \in (t_k, t_{k+1}). \tag{63}$$

Since mismatch control policy w^* is stabilizing (Lewis et al., 2012), it holds that $\|F(\chi) + G(\chi)w^*\| \le \varrho\|\chi\|$ for $\varrho > 0$. By Assumptions 2 and 4, the derivative is upper bounded as

$$\left\|\dot{\bar{e}}_s\right\| \le G_M L_u\|\bar{e}_s\| + L_u\varrho\|\chi\| \le \kappa_1\|\bar{e}_s\| + \kappa_2 \tag{64}$$

where $\kappa_1 = G_M L_u + L_u\varrho$ and $\kappa_3 = L_u\varrho\|\chi_s\|$. Solving the differential inequality by using comparison lemma (Khalil, 1996), the solution is upper bounded by

$$\|\bar{e}_s\| \le \frac{\kappa_2}{\kappa_1}(e^{\kappa_1(t_{k+1}-t_k)} - 1), \quad t \in (t_k, t_{k+1}). \tag{65}$$

Substituting $t = t_{k+1}$ and equating the minimum threshold value r at time t_{k+1} reveals

$$r = \|\bar{e}_s(t_{k+1})\| \leq \frac{\kappa_2}{\kappa_1}(e^{\kappa_1(t_{k+1}-t_k)} - 1). \tag{66}$$

Solving the above inequality for inter-sample time $\tau_k = t_{k+1} - t_k$ we have

$$\tau_k > \frac{1}{\kappa_1} \ln\left(\frac{\kappa_1 r}{\kappa_2} + 1\right) > 0. \tag{67}$$

Therefore, the minimum inter-event time $\tau_m = \inf_{k\in\{0,\mathbb{N}\}}(\tau_k) = \inf_{k\in\{0,\mathbb{N}\}}(t_{k+1} - t_k) > 0$. This completes the proof.

Proof of Theorem 2. The proof is competed using a common Lyapunov candidate function $L : \Re^{2n+l_o} \times \Re_{\geq 0} \to \Re$ for both flow and jump dynamics and presented using two cases.

Case I: Flow dynamics- Consider a continuously differentiable candidate Lyapunov function given by $L = \upsilon_\chi L_\chi + L_W$ where $L_\chi = V_t^*(e_r, t)$ and $L_W = \frac{1}{2}\tilde{W}^T\tilde{W}$ with $\upsilon_\chi = \frac{\varphi_{min}}{\varpi_1}$ where $\varphi_{min} \triangleq \min_{\tau_k,\forall k\in\{0,\mathbb{N}\}}(\lambda_{min}(\varphi(\tau_k)\varphi(\tau_k)^T)) > 0$ is a constant since the regressor vector is persistently exciting.

Considering the first term, the first derivative along the system dynamics can be represented as $\dot{L}_\chi = V_\chi^{*^T}\dot{\chi} = V_\chi^{*^T}(F + Gw + Ge_u)$.

Since the optimal Hamiltonian $\mathcal{H}(\chi, w^*, e_u^*) = 0$, we have $-e_r^TQe_r - w^{*T}Rw^* + \gamma^2 e_u^{*T}e_u^* = V_\chi^{*T}(F + Gw^* + Ge_u^*)$. From (19), we have $G^T(\chi)V_\chi^* = 2\gamma^2 e_u^*$. Adding and subtracting $V_\chi^{*^T}Gw^* + V_\chi^{*^T}Ge_u^*$ in the first derivative and using Young's inequality to separate cross terms in e_u^* and e_u, the first derivative

$$\dot{L}_\chi = -e_r^TQe_r - w^{*T}Rw^* + V_\chi^{*^T}G(w - w^*) + \gamma^2 e_u^Te_u. \tag{68}$$

Using (31) and (33), $w - w^* = \frac{1}{2}R^{-1}G^T(\chi)\nabla\varepsilon(\chi) + \frac{1}{2}R^{-1}G^T(\chi)\nabla\phi^T(\chi)\tilde{W}$. Substituting the above relation and enforcing the event-sampling condition (42) with the NN approximation of the value function V^*, the first derivative leads to

$$\dot{L}_\chi \leq -\lambda_{min}(Q)\|e_r\|^2 + \frac{1}{2}W^T\nabla\phi g_r\nabla\varepsilon + \frac{1}{2}\nabla\varepsilon^T g_r\nabla\varepsilon + \frac{1}{2}W^Tg_\phi\tilde{W} + \frac{1}{2}\nabla\varepsilon^T g_r\nabla\phi^T\tilde{W} + \frac{1}{4\gamma^2}\hat{W}^Tg_\gamma\hat{W}.$$

Using the bounds ι_1, ι_2, ι_3, ι_4, and ι_5 defined in (47) and applying Young's inequality , the first derivative with the definition $\hat{W} = W - \tilde{W}$ becomes

$$\dot{L}_\chi \leq -\lambda_{min}(Q)\|e_r\|^2 + \varpi_1\|\tilde{W}\|^2 + \varpi_2. \tag{69}$$

Next, considering the second term, the first derivative $\dot{L}_W(t) = \tilde{W}^T\dot{\tilde{W}}$ along the weight estimation error dynamics

$$\begin{aligned}\dot{L}_W(t) &= -\alpha_1\frac{\tilde{W}^T\Delta\phi(\tau_{k-1})\Delta\phi^T(\tau_{k-1})\tilde{W}}{(1 + \Delta\phi^T(\tau_{k-1})\Delta\phi(\tau_{k-1}))^2} \\ &\quad -\alpha_1\frac{\tilde{W}^T\Delta\phi(\tau_{k-1})\Delta\varepsilon(\tau_{k-1})}{(1 + \Delta\phi^T(\tau_{k-1})\Delta\phi(\tau_{k-1}))^2} \\ &= -\alpha_1\tilde{W}^T\varphi\varphi^T\tilde{W} - \alpha_1\frac{\tilde{W}^T\varphi\Delta\varepsilon(\tau_{k-1})}{(1 + \Delta\phi^T(\tau_{k-1})\Delta\phi(\tau_{k-1}))^2}\end{aligned} \tag{70}$$

where $\varphi = \varphi(\tau_{k-1})$ is defined in Assumption 7. Applying Young's inequality and separating the cross terms $\tilde{W}$ and ε, the first derivative

$$\dot{L}_W(t) \leq -\alpha_1\varphi_{min}\|\tilde{W}\|^2 + \frac{\alpha_1}{4}\Delta\varepsilon_M^2. \tag{71}$$

Finally, using the first derivatives (69) and (71), the overall first derivative during the flow interval $\dot{L}$ becomes

$$\dot{L} = \upsilon_\chi\dot{L}_\chi + \dot{L}_W \leq -\upsilon_\chi\lambda_{min}(Q)\|e_r\|^2 - \varphi_{min}(\alpha_1 - 1)\|\tilde{W}\|^2 + \varpi_3 \tag{72}$$

where ϖ_3 is defined earlier and learning gain $0 < \alpha_1 < 1$. Define $\varpi_4 = \min(\upsilon_\chi\lambda_{min}(Q), \varphi_{min}(\alpha_1 - 1))$. From (72), the Lyapunov first derivative during the flow period $\dot{L} \leq 0$ as long as $\|e_r\| > \sqrt{\varpi_3/\varpi_4} \triangleq \mathcal{B}_{r_1}$ or $\|\tilde{W}\| > \mathcal{B}_{r_1}$ Therefore, the tracking error e_r and the NN weight estimation error $\tilde{W}$ are ultimately bounded (Haddad et al., 0000). It only remains to show the boundedness at the jump instants to complete the proof.

Case II: Jump dynamics — Considering the same Lyapunov function as in the flow case, the first difference can be expressed as

$$\Delta L = \nu_\chi V^{*+}(e_r, t) - \nu_\chi V^*(e_r, t) + \frac{1}{2}\tilde{W}^{+T}\tilde{W}^+ - \frac{1}{2}\tilde{W}^T\tilde{W}. \tag{73}$$

Along the jump dynamics (45), the first term ΔL_χ in (73) can be expressed as

$$\Delta L_\chi = \nu_\chi V^*(e_r, t) - \nu_\chi V^*(e_r, t) = 0. \tag{74}$$

The first difference of the second term $\Delta L_W = \frac{1}{2}\tilde{W}^{+T}\tilde{W}^+ - \frac{1}{2}\tilde{W}^T\tilde{W}$ along the parameter error dynamics for the jump instants in (45),

$$\begin{aligned}\Delta L_W &= \frac{1}{2}\left[\tilde{W} - \alpha_2\frac{\Delta\phi(\tau_{k-1})\Delta\phi^T(\tau_{k-1})\tilde{W}}{(1 + \Delta\phi^T(\tau_{k-1})\Delta\phi(\tau_{k-1}))^2}\right. \\ &\quad \left. -\ \alpha_2\frac{\Delta\phi(\tau_{k-1})\Delta\varepsilon(\tau_{k-1})}{(1 + \Delta\phi^T(\tau_{k-1})\Delta\phi(\tau_{k-1}))^2}\right]^T \\ &\quad \times \left[\tilde{W} - \alpha_2\frac{\Delta\phi(\tau_{k-1})\Delta\phi^T(\tau_{k-1})\tilde{W}}{(1 + \Delta\phi^T(\tau_{k-1})\Delta\phi(\tau_{k-1}))^2}\right. \\ &\quad \left. -\ \alpha_2\frac{\Delta\phi(\tau_{k-1})\Delta\varepsilon(\tau_{k-1})}{(1 + \Delta\phi^T(\tau_{k-1})\Delta\phi(\tau_{k-1}))^2}\right] \\ &\quad -\ \frac{1}{2}\tilde{W}^T\tilde{W}.\end{aligned} \tag{75}$$

Expanding the first difference and separating the cross terms by applying Young's inequality, it reveals that

$$\begin{aligned}\Delta L_W \leq \frac{1}{2}&\left[-2\alpha_2\frac{\tilde{W}^T\Delta\phi(\tau_{k-1})\Delta\phi^T(\tau_{k-1})\tilde{W}}{(1 + \Delta\phi^T(\tau_{k-1})\Delta\phi(\tau_{k-1}))^2}\right. \\ &+\alpha_2\frac{\tilde{W}^T\Delta\phi(\tau_{k-1})\Delta\phi^T(\tau_{k-1})\tilde{W}}{(1 + \Delta\phi^T(\tau_{k-1})\Delta\phi(\tau_{k-1}))^2} \\ &+\alpha_2\frac{\Delta\varepsilon^T(\tau_{k-1})\Delta\varepsilon(\tau_{k-1})}{(1 + \Delta\phi^T(\tau_{k-1})\Delta\phi(\tau_{k-1}))^2} \\ &+\alpha_2^2\frac{\tilde{W}^T\Delta\phi(\tau_{k-1})\Delta\phi^T(\tau_{k-1})\Delta\phi(\tau_{k-1})\Delta\phi^T(\tau_{k-1})\tilde{W}}{(1 + \Delta\phi^T(\tau_{k-1})\Delta\phi(\tau_{k-1}))^4} \\ &+\alpha_2^2\frac{\tilde{W}^T\Delta\phi(\tau_{k-1})\Delta\phi^T(\tau_{k-1})\Delta\phi(\tau_{k-1})\Delta\phi^T(\tau_{k-1})\tilde{W}}{(1 + \Delta\phi^T(\tau_{k-1})\Delta\phi(\tau_{k-1}))^4} \\ &+\alpha_2^2\frac{\Delta\varepsilon^T(\tau_{k-1})\Delta\phi(\tau_{k-1})\Delta\varepsilon(\tau_{k-1})}{(1 + \Delta\phi^T(\tau_{k-1})\Delta\phi(\tau_{k-1}))^4} \\ &\left.+\alpha_2^2\frac{\Delta\varepsilon^T(\tau_{k-1})\Delta\phi^T(\tau_{k-1})\Delta\phi(\tau_{k-1})\Delta\varepsilon(\tau_{k-1})}{(1 + \Delta\phi^T(\tau_{k-1})\Delta\phi(\tau_{k-1}))^4}\right].\end{aligned} \tag{76}$$

Combining similar terms the first derivative leads to

$$\begin{aligned} \Delta L_W \; \leq \; & \frac{1}{2}\Bigg[-\alpha_2 \frac{\tilde{W}^T \Delta\phi(\tau_{k-1})\Delta\phi^T(\tau_{k-1})\tilde{W}}{(1+\Delta\phi^T(\tau_{k-1})\Delta\phi(\tau_{k-1}))^2} \\ & + 2\alpha_2^2 \frac{\tilde{W}^T \Delta\phi(\tau_{k-1})\Delta\phi^T(\tau_{k-1})\Delta\phi(\tau_{k-1})\Delta\phi^T(\tau_{k-1})\tilde{W}}{(1+\Delta\phi^T(\tau_{k-1})\Delta\phi(\tau_{k-1}))^4} \\ & + \alpha_2^2 \frac{\Delta\varepsilon^T(\tau_{k-1})\Delta\phi(\tau_{k-1})\Delta\varepsilon(\tau_{k-1})}{(1+\Delta\phi^T(\tau_{k-1})\Delta\phi(\tau_{k-1}))^4} \\ & + \alpha_2 \frac{\Delta\varepsilon^T(\tau_{k-1})\Delta\varepsilon(\tau_{k-1})}{(1+\Delta\phi^T(\tau_{k-1})\Delta\phi(\tau_{k-1}))^2} \\ & + \alpha_2^2 \frac{\Delta\varepsilon^T(\tau_{k-1})\Delta\phi^T(\tau_{k-1})\Delta\phi(\tau_{k-1})\Delta\varepsilon(\tau_{k-1})}{(1+\Delta\phi^T(\tau_{k-1})\Delta\phi(\tau_{k-1}))^4} \Bigg]. \end{aligned} \tag{77}$$

With Frobenius norm and the upper bounds for the normalized regression vectors in (48), the first difference is upper bounded as

$$\begin{aligned} \Delta L_W \; \leq \; & -\alpha_2 \varphi_{min}(\frac{1}{2}-\alpha_2)\|\tilde{W}\|^2 + \frac{1}{2}\alpha_2^2 \Delta\varepsilon_M^2 + \frac{1}{2}\alpha_2 \Delta\varepsilon_M^2 \\ & + \frac{1}{2}\alpha_2^2 \Delta\phi_M \Delta\varepsilon_M^2. \end{aligned} \tag{78}$$

Finally, combining both the first differences (74) and (78), the overall first difference at the jump instants

$$\Delta L \leq -\alpha_2 \varphi_{min}(\frac{1}{2}-\alpha_2)\|\tilde{W}\|^2 + \varpi_5 \tag{79}$$

where the learning gain $0 < \alpha_2 < \frac{1}{2}$. From (79), the first difference of the Lyapunov function is negative as long as $\|\tilde{W}\| > \sqrt{\frac{\varpi_5}{\alpha_2 \varphi_{min}(\frac{1}{2}-\alpha_2)}} \triangleq \mathcal{B}_{r_2}$. Therefore, there exists an integer $N > 0$ such that for all $t_k > t_N$, the NN weight estimation error converges to the ultimate bound (Haddad et al., 0000).

From both the flow interval in Case I and jump instants in Case II, the tracking error e_r and the NN weight estimation error $\tilde{W}$ remain bounded both during the flow and at the jump instants and converge a ball of radius $\mathcal{B}_r = \max(\mathcal{B}_{r_1}, \mathcal{B}_{r_2})$ for all $t_k > t_N$. Further, $\|V^* - \hat{V}\|$ and $\|w_s - w_s^*\|$ are ultimately bounded since e_r and $\tilde{W}$ converge to the ultimate bound $\mathcal{B}_r$.

References

Astrom, K. J., & Wittenmark, B. (2013). *Computer-controlled systems: theory and design*. Courier Corporation.
Başar, T., & Bernhard, P. (1995). *H_∞- optimal control and related minimax design problems*. Boston: Birkhäuser.
Bellman, R. (2013). *Dynamic programming*. Courier Corporation.
Bertsekas, D. P. (2017). Value and policy iterations in optimal control and adaptive dynamic programming. *IEEE transactions on neural networks and learning systems*, *28*(3), 500–509.
Bertsekas, D. P., Bertsekas, D. P., Bertsekas, D. P., & Bertsekas, D. P. (1995). *Dynamic programming and optimal control, vol. 1*. Athena scientific Belmont, MA.
Cheng, Y., & Ugrinovskii, V. (2016). Event-triggered leader-following tracking control for multivariable multi-agent systems. *Automatica*, *70*, 204–210.
Dierks, T., & Jagannathan, S. (2009). Optimal tracking control of affine nonlinear discrete-time systems with unknown internal dynamics. In *Decision and control, 2009 held jointly with the 2009 28th Chinese control conference. CDC/CCC 2009. Proceedings of the 48th IEEE conference on* (pp. 6750–6755). IEEE.
Gao, H., & Chen, T. (2008). Network-based h_∞ output tracking control. *IEEE Transactions on Automatic control*, *53*(3), 655–667.
Gommans, T., Antunes, D., Donkers, T., Tabuada, P., & Heemels, M. (2014). Self-triggered linear quadratic control. *Automatica*, *50*(4), 1279–1287.
Haddad, W. M., Chellaboina, V., & Nersesov, S. G. (0000). Impulsive and hybrid dynamical systems, Princeton Series in Applied Mathematics.
Han, Y., Lu, W., & Chen, T. (2015). Consensus analysis of networks with time-varying topology and event-triggered diffusions. *Neural Networks*, *71*, 196–203.
Heemels, W., & Donkers, M. (2013). Model-based periodic event-triggered control for linear systems. *Automatica*, *49*(3), 698–711.
Heydari, A. (0000). Stability analysis of optimal adaptive control under value iteration us- ing a stabilizing initial policy, IEEE Transactions on Neural Networks and Learn- ing Systems.
Heydari, A., & Balakrishnan, S. N. (2014). Fixed-final-time optimal tracking control of input-affine nonlinear systems. *Neurocomputing*, *129*, 528–539.
Ioannou, P., & Fidan, B. (2006). *Advances in design and control*: *vol. 1*, *Adaptive control tutorial*. PA: SIAM.
Kamalapurkar, R., Dinh, H., Bhasin, S., & Dixon, W. E. (2015). Approximate optimal trajectory tracking for continuous-time nonlinear systems. *Automatica*, *51*, 40–48.
Khalil, H. K. (1996). *Nonlinear systems, vol. 2(5)*. New Jersey: Prentice-Hall, 5–1.
Lewis, F., Jagannathan, S., & Yesildirak, A. (1998). *Neural network control of robot manipulators and non-linear systems*. CRC Press.
Lewis, F. L., Vrabie, D., & Syrmos, V. L. (2012). *Optimal control*. John Wiley & Sons.
Liu, J., Tang, J., & Fei, S. (2016). Event-triggered h_∞ filter design for delayed neural network with quantization. *Neural Networks*, *82*, 39–48.
Mazo, M., Jr., Anta, A., & Tabuada, P. (2010). An ISS self-triggered implementation of linear controllers. *Automatica*, *46*(8), 1310–1314.
Modares, H., & Lewis, F. L. (2014). Optimal tracking control of nonlinear partially-unknown constrained-input systems using integral reinforcement learning. *Automatica*, *50*(7), 1780–1792.
Mu, C., Sun, C., Song, A., & Yu, H. (2016). Iterative GDHP-based approximate optimal tracking control for a class of discrete-time nonlinear systems. *Neurocomputing*, *214*, 775–784.
Narayanan, V., & Jagannathan, S. (2016). Approximate optimal distributed control of uncertain nonlinear interconnected systems with event-sampled feedback. In *Decision and control (CDC), 2016 IEEE 55th conference on* (pp. 5827–5832). IEEE.
Peng, C., Song, Y., Xie, X. P., Zhao, M., & Fei, M. -R. (2016). Event-triggered output tracking control for wireless networked control systems with communication delays and data dropouts. *IET Control Theory & Applications*, *10*(17), 2195–2203.
Postoyan, R., Bragagnolo, M. C., Galbrun, E., Daafouz, J., Nešić, D., & Castelan, E. B. (2015). Event-triggered tracking control of unicycle mobile robots. *Automatica*, *52*, 302–308.
Sahoo, A., Narayanan, V., & Jagannathan, S. (2017). Optimal sampling and regulation of uncertain interconnected linear continuous time systems. In *Computational intelligence (SSCI), 2017 IEEE symposium series on* (pp. 1–6). IEEE.
Sahoo, A., Xu, H., & Jagannathan, S. (2017). Approximate optimal control of affine nonlinear continuous-time systems using event-sampled neurodynamic programming. *IEEE Transactions on Neural Networks and Learning Systems*, *28*(3), 639–652.
Sutton, R. S., & Barto, A. G. (1998). *Reinforcement learning: An introduction, vol. 1*. MIT press Cambridge.
Tabuada, P. (2007). Event-triggered real-time scheduling of stabilizing control tasks. *IEEE Transactions on Automatic Control*, *52*(9), 1680–1685.
Tallapragada, P., & Chopra, N. (2013). On event triggered tracking for nonlinear systems. *IEEE Transactions on Automatic Control*, *58*(9), 2343–2348.
Vamvoudakis, K. G., & Ferraz, H. (2016). Event-triggered h-infinity control for unknown continuous-time linear systems using q-learning. In *Decision and control (CDC), 2016 IEEE 55th conference on* (pp. 1376–1381). IEEE.
Vamvoudakis, K. G., Mojoodi, A., & Ferraz, H. (2017). Event-triggered optimal tracking control of nonlinear systems. *International Journal of Robust and Nonlinear Control*, *27*(4), 598–619.
Wang, D., & Liu, D. (0000). Neural robust stabilization via event-triggering mechanism and adaptive learning technique, Neural Networks.
Wang, D., Liu, D., Zhang, Y., & Li, H. (2018). Neural network robust tracking control with adaptive critic framework for uncertain nonlinear systems. *Neural Networks*, *97*, 11–18.
Wang, D., Mu, C., Zhang, Q., & Liu, D. (2016). Event-based input-constrained nonlinear h state feedback with adaptive critic and neural implementation. *Neurocomputing*, *214*, 848–856.
Yang, X., & He, H. (2018). Self-learning robust optimal control for continuous-time nonlinear systems with mismatched disturbances. *Neural Networks*, *99*, 19–30.
Zhang, H., Cui, L., Zhang, X., & Luo, Y. (2011). Data-driven robust approximate optimal tracking control for unknown general nonlinear systems using adaptive dynamic programming method. *IEEE Transactions on Neural Networks*, *22*(12), 2226–2236.
Zhang, H., Wei, Q., & Luo, Y. (2008). A novel infinite-time optimal tracking control scheme for a class of discrete-time nonlinear systems via the greedy hdp iteration algorithm. *IEEE Transactions on Systems, Man and Cybernetics, Part B (Cybernetics)*, *38*(4), 937–942.
Zhang, Q., Zhao, D., & Zhu, Y. (2017). Event-triggered h control for continuous-time nonlinear system via concurrent learning. *IEEE Transactions on Systems, Man, Cybernetics*, *47*(7), 1071–1081.
Zhong, X., & He, H. (2017). An event-triggered ADP control approach for continuous-time system with unknown internal states. *IEEE Transactions on Cybernetics*, *47*(3), 683–694.