2018 IEEE Conference on Decision and Control (CDC)
Miami Beach, FL, USA, Dec. 17-19, 2018

Event-based Near Optimal Sampling and Tracking Control of Nonlinear Systems

Avimanyu Sahoo and Vignesh Narayanan

Abstract— **This paper presents a near optimal event-based tracking control scheme for nonlinear continuous time systems. In order to simultaneously design the event-based sampling intervals and the control policy, the problem of designing the event-triggering mechanism and the feedback controller is posed as a min-max optimization problem. Using the resultant saddle point solution, the feedback control policy and the threshold for the event-based sampling condition is designed. The proposed control scheme is realized by approximating the solution to the associated Hamilton-Jacobi-Issac (HJI) equation using event-based neural networks (NN). The NN weights are updated using an impulsive update scheme. Extension of Lyapunov stability analysis for the impulsive hybrid dynamical system is utilized to prove the local ultimate boundedness of the tracking and NN weight estimation errors. Furthermore, Zeno free behavior of the event-triggering mechanism is guaranteed along with the numerical simulation to corroborate the analytical design.**

I. Introduction

Event-based sampling and control [1] - [5] provides an unified approach for implementing controllers on digital platform. In the event-based control schemes, the sampling intervals are determined on a "as needed" basis. This typically depends on the stability and performance requirements while trying to save computational and network resources. Recently, various event-triggered control schemes are presented in the literature for both state regulation [1] and tracking control [2]- [3] problems. The main idea behind the emulation based event-triggered control design is the design of triggering condition to determine the sampling instants such that the event-based implementation of the controller guarantees stability.

Optimal control [6]- [14] approaches, on the other hand, are employed to optimize the cumulative performance cost in addition to the basic requirement of stability. To design an optimal control for nonlinear continuous time systems, it is required to solve the Hamilton-Jacobi-Bellman (HJB) equation [7]. Adaptive dynamic programming (ADP) and reinforcement learning (RL) [7], [9] techniques are used to obtain an approximate solution for the HJB equation due to the difficulty in obtaining a closed-form solution [8], [9]. A great deal of research results on ADP/RL based optimal control for both regulation [10] and tracking [11]- [12] problems are presented in the literature. These traditional ADP/RL schemes use continuous/periodic state feedback for execution and are computational intensive.

Furthermore, the optimal control problems using ADP/RL are also studied in the context of event-based sampling due to its efficacy in the networked control architectures. In general, the optimal control problem in an event-based framework is required to address two optimization challenges: optimization of the control policy and optimization of triggering instants. However, the recent works on event-based optimal state regulation [4]- [5] and tracking [3] schemes considered the optimization of control policy only. A triggering condition is designed to retain stability. In our recent work, [13], the co-optimization of sampling interval and control policy is presented for regulation of linear systems.

In summary, a co-optimization for tracking control of nonlinear systems is not attempted in the literature. Therefore, in this paper, an event-based co-optimal trajectory tracking control is presented for nonlinear continuous-time systems. The closest works in the literature are by the authors in [11] and [3]. However, the work [11] uses continuous feedback to solve the time-invariant HJB equation whereas in [3] the event-based optimal tracking design only optimizes the control policy by solving the corresponding HJB equation of the discounted cost function. In this work, we propose a novel performance index for the co-optimization problem which leads to the Hamilton-Jacobi-Issac (HJI) equation and a completely different problem.

The event-based tracking error system is formulated by incorporating the error between the continuous and the event-based control policy, referred to as *sampled error policy*, in addition to the tracking error. A system transformation, similar to [11], but, in an event-based formalism, is proposed to avoid the difficulties in the time-varying nature of the value function. The optimization problem is formulated as a min-max problem by introducing a novel time-invariant performance index. The saddle point solution for this problem [14] leads to the maximization of the event-based sampling intervals and minimization of the control policy, simultaneously. Approximate solution to the HJI equation is obtained using a single critic neural network (NN).

The main contributions of the paper include: 1) a novel optimal event-based tracking control scheme to simultaneously optimize the sampling intervals and control policy with the help of a novel performance index; 2) development of an event-based sampling condition with worst case sampled error as threshold with Zeno free behavior; 3) development of the impulsive on-line NN learning scheme; 4) stability analysis of the closed-loop system using extension for impulsive hybrid system.

The remaining of the paper is organized as follows. Section II presents a brief background on the traditional tracking control problem and formulates the problem. In Section

This research is funded start-up fund Oklahoma State University. Email: (avimanyu.sahoo@okstate.edu and vnxv4@mst.edu).

III, a solution to the simultaneous optimization problem is presented. An approximate solution using ADP is presented in Section IV followed by the simulation results in Section V and conclusions in Section VI.

II. Background and Problem Statement

A. Background

Consider a nonlinear continuous time system in an input affine form given by

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

where $x(t) \in \Omega_x \subset \mathbb{R}^n$ and $u(t) \in \mathbb{R}^m$ are, respectively, the state and the control input; Ω_x is a compact set in the n-dimensional Euclidean space. The nonlinear functions $f : \Omega_x \to \mathbb{R}^n$ and $g : \Omega_x \to \mathbb{R}^{n\times m}$ are the internal dynamics and the input gain, respectively, with $f(0) = 0$.

The control objective for the system state $x(t)$ is to track a desired state trajectory $x_d(t) \in \mathbb{R}^n$. The trajectory is generated by a reference system given by

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

where $\zeta : \Omega_{x_d} \to \mathbb{R}^n$ is the internal dynamics, $x_d(t) \in \mathbb{R}^n$ is the desired state with $\zeta(0) = 0$.

Assumption 1: The system (1) is stabilizable and the system states are available for measurement. The functions $f(x(t))$ and $g(x(t))$ are Lipschitz continuous in Ω_x. The function $g(x(t))$ is full column rank and satisfies $\|g(x(t))\| \le g_M$ for some $g_M > 0$ and $g(x_d)g^+(x_d) = I$ where $g^+ = (g^Tg)^{-1}g^T$.

Assumption 2: The reference trajectory $x_d(t)$ is bounded such that $\|x_d(t)\| \le b_{x_d} \in \mathbb{R}$.

To ensure that the trajectory tracking objective is tractable, define the tracking error $e_r(t) \triangleq x(t) - x_d(t)$. The dynamics of the tracking error, from (1) and (2), can be written as

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

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 steady-state feedback control policy [11] can be expressed as

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

where u_d is the expected control policy corresponding to the desired trajectory.

Define an augmented tracking error system state $z \triangleq [e_r^T \;\; x_d^T]^T \in \mathbb{R}^{2n}$. The dynamics of the augmented system can be represented as

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

where the nonlinear functions $F : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ is given by $F(z) \triangleq \begin{bmatrix} f(e_r + x_d) + g(e_r + x_d)u_d - \zeta(x_d) \\ \zeta(x_d) \end{bmatrix}$, $G : \mathbb{R}^{2n} \to \mathbb{R}^{2n\times m}$ given by $G(z) \triangleq \begin{bmatrix} g(e_r + x_d) \\ 0 \end{bmatrix}$, and the mismatch control policy $w \triangleq u - u_d \in \mathbb{R}^m$. Note that $f(0) = 0$ implies the augmented $F(0) = 0$.

Define the cost functional subject to the dynamical constraint in (5) as

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

where $\bar{Q} \triangleq \begin{bmatrix} Q & 0_{n\times n} \\ 0_{n\times n} & 0_{n\times n} \end{bmatrix} \in \mathbb{R}^{2n\times 2n}$ where $Q \in \mathbb{R}^{n\times n}$ is a positive definite and $R \in \mathbb{R}^{m\times m}$ is symmetric positive definite matrix. The matrices $0_{n\times n}$ are matrices with all elements zero. Since the augmented dynamics in (5) transfered the time-varying problem to a time-invariant problem [11], the cost functional is finite for any admissible control policy.

An approximate solution of the optimal tracking control problem can be obtained by approximating the solution of the corresponding HJB equation using NN approximation [11] with continuous/periodic availability of the state feedback information. The optimal tracking control problem for event-based control is formulated next.

B. Problem Definition

In an event-triggered control formalism the system state is sampled aperiodically. Define the aperiodic sampling instants as a sequence $\{t_k\}_{k\in\{0,\mathbb{N}\}} \subseteq \{t\}$, such that $0 = t_0 < t_1 < \cdots$. The system states are sampled and the control policies are updated at the instant t_k, $k = 0, 1, \cdots$.

Define the piecewise constant event-sampled state at the controller as $x_s(t) = x(t_k), \; \forall t \in [t_k, t_{k+1})$. With the event-based availability of the state information $x_s(t)$, the control policy $u_s(t)$is held at the actuator using a zero-order hold till the next update and a piecewise constant function.

The nonlinear system given by (1) with event-based control policy $u_s(t)$ can be expressed as

$$\dot{x} = f(x) + g(x)u_s. \tag{7}$$

The error between the continuous control input $u(t)$ and the event-sampled control input $u_s(t)$, referred to as sampled error policy, $e_u(t) \in \mathbb{R}^m$, is given by

$$e_u(t) = u_s(t) - u(t) \tag{8}$$

Using (8), the system in (7) can be represented as

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

The tracking error dynamics, with the event-triggered system (9) and the desired state trajectory (2) can be expressed as

$$\dot{e}_r(t) = f(e_r + x_d) + g(e_r + x_d)u + g(e_r + x_d)e_u - \zeta(x_d). \tag{10}$$

Then, the augmented system can be represented as

$$\dot{z} = F(z) + G(z)w + G(z)e_u. \tag{11}$$

The problem in hand is to design the threshold for the event-triggering mechanism, which will maximize a performance index, and the control policy, to minimize a performance index. Therefore, the time-invariant performance index in (6) needs to be redefined to account for the threshold for the sampled error policy. A solution to the above problem is presented next.

III. Performance Index and Co-optimization

In this section, a min-max optimization problem is proposed and the saddle-point solution to the min-max co-optimization problem is obtained.

Define the event-based augmented state available at the controller as

$$z_s(t) = z(t_k),\ \forall t \in [t_k, t_{k+1}) \tag{12}$$

where $z(t_k) = [e_r^T(t_k), x_d^T(t_k)]^T$. The error between the continuous augmented state $z(t)$ and the sampled augmented state $z_s(t)$, referred to as state sampling error, is given by

$$e_s(t) = z_s(t) - z(t), \qquad \forall t \in [t_k, t_{k+1}). \tag{13}$$

The sampling intervals in the event-based control scheme can be maximized by allowing the sampled error policy (8) to increase to its limiting value (maximum threshold value) without compromising the system stability before the next sampling.

To obtain the worst case threshold of the sampled error policy, we redefine the cost functional (6) as

$$\begin{aligned} J(z, w, \hat{e}_u) = \int_0^\infty & [z^T(\tau)\bar{Q}z(\tau) + w(\tau)^T R w(\tau) \\ & - \gamma^2 \hat{e}_u^T(\tau)\hat{e}_u(\tau)] d\tau \end{aligned} \tag{14}$$

where $\gamma > \gamma^*$ represents the attenuation constant [14] and $\hat{e}_u$ is the threshold for sampled error policy e_u.

The optimization problem is posed as a minimax problem where the mismatch control policy, w, is the minimizing player and the threshold for sampled error policy, $\hat{e}_u$, is the maximizing player. The objective is the find the optimal mismatch control policy, w^*, and worst case threshold, e_u^* for sampled error policy e_u such that the saddle-point solution $\underset{w}{min}\ \underset{\hat{e}_u}{max}\ J(z, w, \hat{e}_u) = \underset{\hat{e}_u}{max}\ \underset{w}{min}\ J(z, w, \hat{e}_u)$ is reached. The saddle-point optimal value function, $V^* : \mathbb{R}^{2n} \to \mathbb{R}_{\geq 0}$ can be expressed as

$$V^*(z(t)) = \underset{w(\tau)|\tau \in \mathbb{R}_{\geq t}}{min}\ \underset{\hat{e}_u(\tau)|\tau \in \mathbb{R}_{\geq t}}{max} J(z, w, \hat{e}_u). \tag{15}$$

The Hamiltonian, with the admissible control policy and state constraint (11), is defined as

$$\begin{aligned} \mathcal{H}(z, w, \hat{e}_u) = z^T\bar{Q}z + w^T R w - \gamma^2 \hat{e}_u^T \hat{e}_u \\ + V_z^{*T}[F + Gw + Ge_u] \end{aligned} \tag{16}$$

where $V_z^* = \partial V^*/\partial z$, $V^*(z)$ is the optimal value defined in (15). The optimal mismatch control policy in a closed form, by using $\frac{\partial \mathcal{H}(z,w,\hat{e}_u)}{\partial w} = 0$, is given by

$$w^*(z) = -(1/2)R^{-1}G^T(z)V_z^*(z). \tag{17}$$

Similarly, the worst-case threshold vale for the sampled error policy in a closed form, with $\frac{\partial \mathcal{H}(z,w,\hat{e}_u)}{\partial \hat{e}_u} = 0$, is given by

$$e_u^*(z) = (1/2\gamma^2)G^T(z)V_z^*(z). \tag{18}$$

The optimal control policy, u^*, is given by

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

The event-sampled optimal control policy u_s^* is given by

$$u_s^* = -\frac{1}{2}R^{-1}G^T(z_s)V_{z_s}^* + g^+(x_{ds})(\zeta(x_{ds}) - f(x_{ds})). \tag{20}$$

where $V_{z_s}^* = \frac{\partial V^*(z)}{\partial z}|_{z=z_s}$. The HJI equation with optimal mismatch policies (17) and (18) is given by

$$\begin{aligned} &\mathcal{H}^* = z^T\bar{Q}z + w^{*T}Rw^* - \gamma^2 e_u^{*T}e_u^* + V_z^{*T}[F + \\ &Gw^* + Ge_u^*] = 0 \end{aligned} \tag{21}$$

for all z with $V^*(0) = 0$. The existence of the solution of the HJI equation (21), i.e., the optimal value function, is guaranteed for a reachable and zero-state observable system with $\gamma > \gamma^*$, where γ^* is the H_∞ gain [14].

As the closed-form solution to the HJI equation is almost impossible to compute analytically [8], an approximate solution is presented next.

IV. Approximation of the Optimal Solution

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

Assuming that the solution to the HJI equation, i.e., the optimal value function $V^* : \mathbb{R}^{2n} \to \mathbb{R}_{\geq 0}$, exists and continuously differentiable the value function using a NN can be represented as

$$V^*(z) = W^T\phi(z) + \varepsilon(z) \tag{22}$$

where $W \in \mathbb{R}^{l_o}$ is the unknown constant target weight vector, $\phi : \mathbb{R}^{2n} \to \mathbb{R}^{l_o}$ is the smooth activation function satisfying $\phi(0) = 0$ with l_o hidden layer neurons. The function $\varepsilon : \mathbb{R}^{2n} \to \mathbb{R}$ is the reconstruction error. The following standard assumption for the NN is used for analysis.

Assumption 3: The unknown target weight vector $W \in \mathbb{R}^{l_o}$ is bounded above and given by $\|W\| \leq W_M$ where $W_M \in \mathbb{R}$ is a positive constant. The activation function ϕ is bounded, continuously differentiable with bound given by $\|\phi(z)\| \leq \phi_M$. Further, the NN reconstruction error is bounded such that $\|\varepsilon(z)\| \leq \varepsilon_M$ and $\|\nabla\varepsilon(z)\| \leq \bar{\varepsilon}_M$ where $\nabla\varepsilon(z) = \frac{\partial \varepsilon(z)}{\partial z}$.

With the NN approximation of the optimal value function, the optimal mismatch policy can be computed as

$$w^*(z) = -(1/2)R^{-1}G^T(z)(\nabla\phi^T(z)W + \nabla\varepsilon(z)), \tag{23}$$

and the worst case threshold of sampled error policy is

$$e_u^*(z) = -(1/2\gamma^2)G^T(z)(\nabla\phi^T(z)W + \nabla\varepsilon(z)), \tag{24}$$

where $\nabla\phi(z) = \frac{\partial \phi(z)}{\partial z}$. Based on (22), the NN approximation with estimated NN weights is given by

$$\hat{V}(z) = \hat{W}^T\phi(z_s), \quad \forall t \in [t_k, t_{k+1}). \tag{25}$$

where $\hat{W} \in \mathbb{R}^{l_o}$ NN estimated weights and $\phi(z_s) \in \mathbb{R}^{l_o}$ is the event-sampled activation function which uses the sampled augmented state as input. Then, estimated mismatch control policy

$$w(z) = -(1/2)R^{-1}G^T(z)(\nabla\phi^T(z_s)\hat{W}), \tag{26}$$

and the estimated control input can be expressed as

$$u = -\frac{1}{2}R^{-1}G^T(z)\nabla\phi^T(z_s)\hat{W} + g^+(x_d)(\zeta(x_d) - f(x_d)). \quad (27)$$

From (27), the event-sampled control input applied to the system (1) with event-based state z_s and (4) can be expressed as

$$\begin{aligned} u_s &= -(1/2)R^{-1}G^T(z_s)\nabla\phi^T(z_s)\hat{W} \\ &+ g^+(x_{ds})(\zeta(x_{ds}) - f(x_{ds})), \forall t \in [t_k, t_{k+1}). \end{aligned} \quad (28)$$

The estimated threshold of sampled error policy with the approximated value function is given by

$$\hat{e}_u = (1/2\gamma^2)G^T(z)\nabla\phi^T(z_s)\hat{W}. \quad (29)$$

In an event-based sampling framework, applying Bellman principle of optimality [8], event-based Bellman equation can be expressed as

$$\begin{aligned} V^*(t_{k+1}) - V^*(t_k) = \int_{t_k}^{t_{k+1}} &\left(-z^T\bar{Q}z - w^TRw \right. \\ &\left. +\gamma^2\hat{e}_u^T\hat{e}_u\right) d\tau. \end{aligned} \quad (30)$$

The Bellman equation (30), by inserting the NN representation of the value function (22), yields

$$\begin{aligned} W^T\Delta\phi(\tau_k) + \Delta\varepsilon(\tau_k) = \textstyle\int_{t_k}^{t_{k+1}}(&-z^T\bar{Q}z - w^TRw \\ &+ \gamma^2\hat{e}_u^T\hat{e}_u)d\tau \end{aligned} \quad (31)$$

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

With the approximated solution of the value function in (25), the Bellman error is given as.

$$\begin{aligned} \delta_{k+1} &= \textstyle\int_{t_k}^{t_{k+1}} (z^T\bar{Q}z + w^TRw - \gamma^2\hat{e}_u^T\hat{e}_u)d\tau \\ &+ \hat{V}(t_{k+1}) - \hat{V}(t_k) = \textstyle\int_{t_k}^{t_{k+1}}(z^T\bar{Q}z \\ &+ w^TRw - \gamma^2\hat{e}_u^T\hat{e}_u)d\tau + \hat{W}^T\Delta\phi(\tau_k) \end{aligned} \quad (32)$$

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

The value function weights are updated to minimize the Bellman residual error. Since the augmented system states are available only at the triggering instants, impulsive weight update laws are selected with both jump and flow dynamics, respectively, given as

$$\hat{W}^+ = \hat{W} - \alpha_2\frac{\Delta\phi(\tau_{k-1})}{(1+\Delta\phi(\tau_{k-1})^T\Delta\phi(\tau_{k-1}))^2}\delta_k^T,\ t = t_k \quad (33)$$

$$\dot{\hat{W}} = -\alpha_1\frac{\Delta\phi(\tau_{k-1})}{(1+\Delta\phi(\tau_{k-1})^T\Delta\phi(\tau_{k-1}))^2}\delta_k^T,\ t \in (t_k, t_{k+1}) \quad (34)$$

where $\alpha_1, \alpha_2 > 0$ are learning gains, $\delta_k = \int_{t_{k-1}}^{t_k}(z^T\bar{Q}z + w^TRw - \gamma^2\hat{e}_u^T\hat{e}_u)d\tau + \hat{W}^T\Delta\phi(\tau_{k-1})$. Note that the integration $\int_{t_{k-1}}^{t_k}(\cdot)d\tau$ and the difference $\Delta\phi(\tau_k)$ uses the state information at two consecutive sampling instants t_k and t_{k-1} which are available at k and, hence, δ_k is in a computable form.

Remark 1: The rationale behind the impulsive parameter update is twofold. First, the availability of the state information aperiodically at the triggering instants leads to a combination of the continuous and discrete dynamics for the closed-loop event-triggered system. Therefore, the NN weights are updated as a jump in the weight with the new state feedback information at the triggering instants $t_k, \forall k \in \{0, \mathbb{N}\}$. Second, the inter-sample times or the flow periods is utilized to update the NN weights with the state information received at the previous triggering instants for a faster convergence, which is motivated by the traditional iteration based techniques ADP schemes [7].

Define the NN weight estimation error as $\tilde{W} = W - \hat{W}$. Then, subtracting (32) from (31), the Bellman error, with a event-step backward, can be represented as

$$\delta_k = -\tilde{W}^T\Delta\phi(\tau_{k-1}) - \Delta\varepsilon(\tau_{k-1}). \quad (35)$$

Note that, Bellman error (35) is expressed as a function of NN weight estimation error and in a unmeasurable form. This is used for demonstrating the stability.

The event-based sampling condition which determines the sampling instants of the closed-loop event-triggered system with the estimated threshold $\hat{e}_u$ in (29) can be defined as

$$\begin{aligned} t_{k+1} = \inf\{t > t_k \mid e_u^Te_u = \max\Bigg(&r^2, \frac{1}{4\gamma^4}\hat{W}^T \\ &\times\nabla\phi(z_s)GG^T(z)\nabla\phi^T(z_s)\hat{W}\Bigg)\}. \end{aligned} \quad (36)$$

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

$$\dot{\xi} = \begin{bmatrix} F(z) + G(z)w + G(z)e_u \\ \alpha_1\frac{\Delta\phi(\tau_{k-1})}{(1+\Delta\phi(\tau_{k-1})^T\Delta\phi(\tau_{k-1}))^2}\delta_k^T \end{bmatrix},\ \xi \in \mathcal{C},\ \ t \in (t_k, t_{k+1}) \quad (37)$$

$$\xi^+ = \begin{bmatrix} z \\ \tilde{W} + \alpha_2\frac{\Delta\phi(\tau_{k-1})}{1+([\Delta\phi(\tau_{k-1})]^T[\Delta\phi(\tau_{k-1})])^2}\delta_k^T \end{bmatrix},\ \xi \in \mathcal{D}\ \ t = t_k. \quad (38)$$

where (37) is the dynamics of the system during the inter-sample time, referred to as flow dynamics, and (38) is the dynamics at the sampling instants referred to as jump dynamics. The sets $\mathcal{C} \triangleq \{\xi \in \mathbb{R}^{2n+l_o} \mid e_u^Te_u < \max(r^2,\ \frac{1}{4\gamma^4}\hat{W}^T\nabla\phi(z)GG^T(z)\nabla^T\phi(z)\hat{W})\}$ and $\mathcal{D} \triangleq \{\xi \in \mathbb{R}^{2n+l_o} \mid e_u^Te_u = \max(r^2,\ \frac{1}{4\gamma^4}\hat{W}^T\nabla\phi(z)GG^T(z)\nabla^T\phi(z)\hat{W})\}$ are the flow and jump sets, respectively.

Theorem 1: Consider the event-sampled augmented tracking error system (11) along with the performance index (14), event-based control policy (28) and NN weight update rule (33) and (34), represented as an nonlinear impulsive hybrid dynamical system (37) and (38) . Suppose the Assumptions 1-3 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 an integer $N > 0$ such that the tracking error and the NN weight estimation error are locally ultimately bounded for all sampling instants $k > N$, provided event-based sampling instants are obtained using (36), the regressor vector satisfies persistency of excitation condition, and the weight tuning gains are selected as $0 < \alpha_1 < 1$ and

$0 < \alpha_2 < \frac{1}{2}\varphi_{min}$. Further, $\|V^* - \hat{V}\|$ and $\|u_s - u_s^*\|$ are also ultimately bounded.

Proof: (sketch) The proof is competed using a common Lypunov candidate function $L : \mathbb{R}^{2n+l_o} \times \mathbb{R}_{\geq 0} \to \mathbb{R}$ for both flow and jump dynamics. The Lyapunov candidate function for flow dynamics is given by $L(\xi, t) = \upsilon_z L_z + L_W$ where $L_z = V_t^*(e_r, t)$ and $L_W = \frac{1}{2}\tilde{W}^T\tilde{W}$ with $\upsilon_z = \frac{\varphi_{min}}{\varpi_1}$ where $\varphi_{min} > 0$ is a constant since the regressor vector is persistently exciting [15].

The first derivative during the flow interval $\dot{L}(\xi, t)$ becomes

$$\begin{aligned}\dot{L}(\xi, t) = \upsilon_z \dot{L}_z + \dot{L}_W \leq -\upsilon_z \lambda_{min}(Q)\|e_r\|^2 \\ -\varphi_{min}(\alpha_1 - 1)\|\tilde{W}\|^2 + \varpi_3\end{aligned} \tag{39}$$

where ϖ_3 is a constantcomputed from the constant bounds, and learning gain $0 < \alpha_1 < 1$. Define $\varpi_4 = \min(\upsilon_z\lambda_{min}(Q), \varphi_{min}(\alpha_1 - 1))$. From (39), the Lyapunov first derivative during the flow period $\dot{L}(\xi, t) \leq 0$ as long as $\|\xi\| > \sqrt{\varpi_3/\varpi_4} \triangleq \mathcal{B}_{\xi_1}$.

The first differences at the jump instants can be expressed as

$$\Delta L \leq -\alpha_2(\tfrac{1}{2}\varphi_{min} - \alpha_2)\|\tilde{W}\|^2 + \varpi_5 \tag{40}$$

where ϖ_5 is a constant, computed from the constant bounds, and the learning gain $0 < \alpha_2 < \frac{1}{2}\varphi_{min}$. From (40), the first difference of the Lyapunov function is negative as long as $\|\xi\| > \sqrt{\frac{\varpi_5}{\alpha_2(\frac{1}{2}\varphi_{min} - \alpha)}} \triangleq \mathcal{B}_{\xi_2}$.

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 $\tilde{W}$ remains bounded both during the flow and at the jump instants and converges a set B_ξ for all $k > N$. Further, $\|V^* - \hat{V}\|$ and $\|u_s - u_s^*\|$ are ultimately bounded since e_r and $\tilde{W}$ converge to the ultimate bound.

Remark 2: Note that the constant r in the triggering condition (36) enforces the minimum inter-event time. Further The bounds are obtained in Theorem 1 as a function of the NN reconstruction error. As the number of hidden layer neurons increase, the reconstruction error converges close to zero.

To show the Zeno free behavior, we will use a conservative triggering condition presented in the next theorem. Before proceeding further, the following assumption and the technical lemma is necessary.

Assumption 4: The optimal policies $u^*(t)$ and $e_u^*(t)$ are locally Lipschitz continuous with respect to $z \in \Omega_z$ and satisfies $\|u_s^* - u^*\| = \|e_u\| \leq L_u \|e_s\|$ and $\|e_u^*(z) - e_u^*(z_s)\| \leq L_u \|e_s\|$ where $L_u > 0$ is the Lipschitz constant.

Lemma 1: The inequality

$$e_u^T e_u \leq (1/4\gamma^4)\hat{V}_{z_s}^T G(z) G^T(z) \hat{V}_{z_s}. \tag{41}$$

holds if the inequality given by

$$L_u\|e_s\| \leq (1/4\gamma^2)\|G^T(z_s)V_{z_s}\| \tag{42}$$

holds.

Proof: The proof is the result of the Lipschitz continuity control policy in Assumption 4.

Assumption 5: For an optimal mismatched policy w^*, the following inequality holds $\|F + Gw^*\| \leq \varrho$ where $\varrho \in \mathbb{R}_{>0}$.

The assumption is trivial since the continuous time optimal control policy is asymptotically stabilizing [6].

Theorem 2: Let the Assumption 1-5 hold. With the event-based sampling condition

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

the tracking error and the NN weight estimation error of event-triggered augmented tracking error system (11), with event-based control policy (28) and NN weight update rule (33) and (34), are ultimately bounded. 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

$$\tau_k = t_{k+1} - t_k > (1/g_M)\ln\left((g_M r/8\varrho\gamma^2) + 1\right). \tag{43}$$

Proof: The proof is omitted due to space limitation. Next, the numerical simulation results and presented.

V. Simulation Results

A two link robot manipulator is considered for the numerical simulation. The dynamics of the robot are given by [11]

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

where $x = [q_1,\ q_2,\ \dot{q}_1,\ \dot{q}_2]^T$ is the state vector, $f = \left[x_3,\ x_4,\ (M^{-1}(-V_m - F_d)\begin{bmatrix}x_3\\x_4\end{bmatrix} - F_s)^T\right]^T$, $g = \left[[0,\ 0]^T,\ [0,\ 0]^T,\ (M^{-1})^T\right]^T$, $M = \begin{bmatrix}p_1 + 2p_3c_2 & p_2 + p_3c_2\\ p_2 + p_3c_2 & p_2\end{bmatrix}$, $V_m = \begin{bmatrix}-p_3s_2\dot{q}_2 & -p_3s_2(\dot{q}_1 + \dot{q}_2)\\ p_3s_2\dot{q}_1 & 0\end{bmatrix}$, $F_d = diag[5.3,\ \ 1.1]$, $p_1 = 3.473, p_2 = 0.196, p_3 = 0.242, c_2 = sin(\dot{q}_2)$, and $s_2 = sin(\dot{q}_2)$.

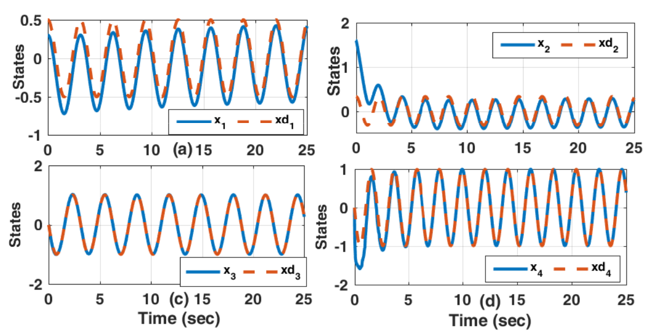


Fig. 1. System state and desired trajectories

The reference trajectory considered for tracking is given by $x_d = [0.5cos(2t),\ 0.33cos(3t),\ -sin(2t),\ -sin(3t)]^T$ with $\zeta(x_d) = [x_{d3},\ x_{d4},\ -4x_{d1},\ -9x_{d2}]^T$. The penalty matrices of performance index (14) were selected as $Q = diag[0.1,\ 1,\ 5,\ 5]$, $R = diag[0.1,\ 0.2]$, $\gamma = 2.5$. The NN weights are initialized randomly from a uniform distribution in the interval $[0,\ 2]$. The learning gains are selected as $\alpha_1 = 0.004$, and $\alpha_2 = 0.26$. A polynomial regression vector as in [11] 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 to ensure the convergence of the NN weights. The simulation is run for 25 secs with initial states $x_0 = [0.3,\ 1.6, 0, 0]^T$.

The time history of the robot states and the desired reference trajectory is shown in Fig. 1. The system states track the desired trajectory and the tracking error converges close to zero as shown in Fig. 2 (a). Further, the Bellman error also converges close to zeros and shown in Fig. 2. (b). This implies the NN approximation of the value function converge to the neighborhood of the optimal value. From the

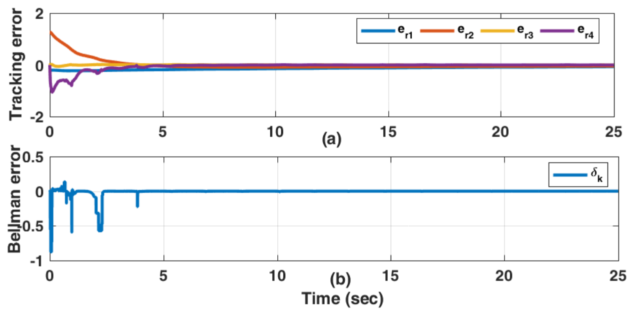
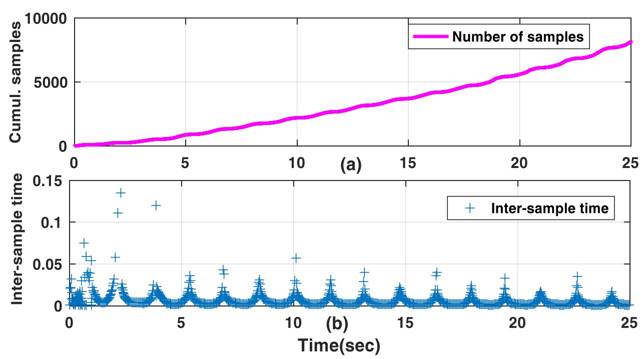


Fig. 2. Convergence of (a) Tracking error (b) event-based integral Bellman error.

event-triggered control perspective, as shown in Fig. 3(a) and 3(b), the event-sampling condition is computed using (36) with $r = 0.001$. It is observed that the control is executed 8138 times for a $\gamma = 2.5$ during the simulation time of 25 sec. This shows a reduction of feedback communication. Note that different value of γ will result in different number of sampling as the penalizing factor for sampled error policy is changed.



inin

Fig. 3. Event-triggered system performance (a) Cumulative number of sampling; (b) inter-sampling instants.

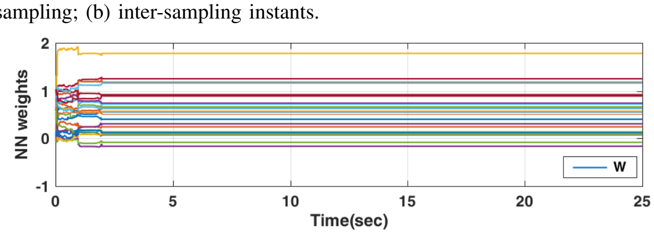


Fig. 4. Convergence of NN weight estimates

The convergence of the estimated NN weights are shown in Fig. 4. Since the target NN weights are unknown, it is not possible to show the convergence of the weight estimates. However, the NN weights become constant after the initial learning which implies the convergence to a close neighborhood of the target values.

VI. Conclusions

An event-based approximate optimal sampling and trajectory tracking control is presented. An approximate solution is obtained using NN based approximation and impulsive learning of the NN weights. The impulsive weight update scheme improved the NN weight convergence and maintained the boundedness of the system during inter-sample times. The triggering condition is designed based on the novel performance index which optimizes the sampling interval along with the control policy. The analytical design is also verified with numerical simulation. The proposed design ensured the convergence of the tracking error and the NN weight estimation error to the ultimate bound. The communication imperfections are not included in the paper for analysis and will be included in our future research.

References

[1] P. Tabuada, "Event-triggered real-time scheduling of stabilizing control tasks," *IEEE Transactions on Automatic Control*, vol. 52, no. 9, pp. 1680-1685, Sep 2007.

[2] P. Tallapragada and N. Chopra, "On event triggered tracking for nonlinear systems," *IEEE Transactions on Automatic Control*, vol. 58, no. 9, pp. 2343–2348, Sept 2013.

[3] K. G. Vamvoudakis, A. Mojoodi, and H. Ferraz, "Event-triggered optimal tracking control of nonlinear systems," Int. J. Robust Nonlinear Control, vol. 27, no. 4, pp. 598619, 2017.

[4] A. Sahoo, H. Xu, and S. Jagannathan, "Approximate optimal control of affine nonlinear continuous-time systems using event-sampled neurodynamic programming," *IEEE Transactions on Neural Networks and Learning Systems*, vol. PP, no. 99, pp. 1–14, 2016.

[5] V. Narayanan and S. Jagannathan, "Approximate optimal distributed control of uncertain nonlinear interconnected systems with event-sampled feedback," in *Proc. of IEEE 55th Conference on Decision and Control (CDC)*, Dec 2016, pp. 5827–5832.

[6] F. L. Lewis and V. L. Syrmos, *Optimal control*. John Wiley & Sons, 1995.

[7] D. P. Bertsekas, *Dynamic programming and optimal control*. vol. 1, no. 2, Athena Scientific Belmont, MA, 1995.

[8] R. Bellman, *Dynamic programming*. Princeton University Press, Princeton, NJ, USA (1957)

[9] R. S. Sutton and A. G. Barto, *Reinforcement learning: An introduction*. MIT press Cambridge, 1998, vol. 1, no. 1.

[10] T. Dierks and S. Jagannathan, "Optimal control of affine nonlinear continuous-time systems," in *Proc. of American Control Conference*, Baltimore, USA, Jun 2010: 1568-1573.

[11] R. Kamalapurkar, H. Dinh, S. Bhasin, and W. E. Dixon, "Approximate optimal trajectory tracking for continuous-time nonlinear systems," Automatica, vol. 51, pp. 4048, 2015.

[12] H. Zhang, Q. Wei, and Y. Luo, "A novel infinite-time optimal tracking control scheme for a class of discrete-time nonlinear systems via the greedy HDP iteration algorithm," *IEEE Trans. Syst. Man, Cybern. Part B Cybern.*, vol. 38, no. 4, pp. 937942, 2008.

[13] A. Sahoo, V. Narayanan and S. Jagannathan, "Optimal sampling and regulation of uncertain interconnected linear continuous time systems," in *Proc.2017 of the IEEE Symposium Series on Computational Intelligence (SSCI)*, Honolulu, HI, 2017, pp. 1-6.

[14] T. Basar and P. Bernhard, *H_∞ optimal control and related minimax design problems*. Birkhaüser, Boston, Massachusetts, 1991.

[15] P. Ioannou, B. Fidan, *Adaptive Control Tutorial.* Advances in Design and Control, PA: SIAM, 2006.

[16] W. M. Haddad, V. Chellaboina, S. G. Nersesov *Impulsive and Hybrid Dynamical Systems: Stability, Dissipativity, and Control.* Princeton, NJ: Princeton University Press, 2006.

[17] H. Khalil, *Nonlinear Systems*. Pearson Education. Prentice Hall, 2002.