2014 International Joint Conference on Neural Networks (IJCNN)
July 6-11, 2014, Beijing, China

Near Optimal Event-Based Control of Nonlinear Discrete Time Systems in Affine Form with Measured Input and Output Data

Avimanyu Sahoo
Department of Electrical and Computer Engineering
Missouri University of S&T
Rolla, MO, USA
asww6@mst.edu

Hao Xu
Department of Engineering
The University of Tennessee at Martin
Martin, TN, USA
hxu12@utm.edu

S. Jagannathan
Department of Electrical and Computer Engineering
Missouri University of S&T
Rolla, MO, USA
sarangap@mst.edu

***Abstract*— In this paper, an event-based near optimal control of uncertain nonlinear discrete time systems is presented by using input-output data and approximate dynamic programming (ADP). The nonlinear system dynamics in affine form are transformed into an input-output form. Then, three neural networks (NN) with event sampled input-output vector are used, namely, the identifier NN to relax the knowledge of the system dynamics, a critic NN to approximate the value function which is the solution to the Hamilton-Jacobi Bellman (HJB) equation, and an actor NN to approximate the optimal control policy, in an online manner without utilizing value or policy iterations. In addition, the NN weights of all the three NNs are tuned only at event-triggered instants leading to a novel non-periodic update rule to reduce computation when compared to traditional NN based scheme. Further, an event-trigger condition to decide the trigger instants is derived. Finally, the Lyapunov technique is used in conjunction with the event-trigger condition to guarantee the uniform ultimate boundedness (UUB) of the closed-loop system. The analytical design is substantiated with numerical results via simulation.**

Keywords*— *Approximate dynamic programming; event-triggered control; Hamilton-Jacobi-Bellman equation; neural networks; Optimal control

I. Introduction

Event-based sampling [2]-[5] of state or output vector for controller execution is gaining popularity among control researchers in the recent time since this control scheme referred to as event-triggered control can save computation. Further, this alternate control paradigm can also save network resources in terms of bandwidth of the communication network in case of the networked control systems (NCS). The sampling or the transmission instants, referred to as event-trigger instants, are decided by a state or output dependent criteria referred to as event-trigger condition.

Recently, various event-triggered control (ETC) schemes [2]-[5] have been introduced in the literature with complete or partial knowledge of the system dynamics using both state feedback [2], [4]-[5] and output feedback [3]. Further, the traditional optimal control design [7] approach is extended to event-triggered control [4] of linear systems by using backward-in-time solution of the Riccati equation (RE) with completely known system dynamics. In contrast, various online approaches using approximate dynamic programming (ADP) [8] are available to solve the optimal control in a forward-in-time manner provided a significant number of iterations within a sampling interval can be utilized which is a major bottleneck [9]. On the other hand, developing an event-triggered optimal controller computed forward in time and online manner without iterations is still an open problem.

Motivated by the above fact, in this paper, we propose for the first time a novel technique to solve the infinite horizon optimal control of nonlinear discrete-time system in an event-triggered paradigm by using measured input-output data and approximate dynamic programming (ADP) approach. The value function and the optimal control input are approximated by using neural networks (NN) referred as actor-critic NNs. Further, to relax the system dynamics, a NN based identifier is used in conjunction with the actor-critic NNs. A novel event-trigger condition is developed which not only reduces the number of controller updates but initiates sufficient number of events during the initial learning phase for NN approximation. Further, to reduce computation when compared to a traditional NN based scheme [6], novel aperiodic update laws are developed which tunes the identifier, actor and critic NN weights once every triggered instant.

The main contributions of this paper include the design of an optimal event-triggered control scheme for an uncertain nonlinear discrete time system in affine form by using input-output event sampled data with a novel event-trigger condition and aperiodic update laws to save computation. Finally, Lyapunov direct method is used to prove the uniform ultimate boundedness (UUB) of the closed-loop event-triggered system. Next, after a brief background the problem for the optimal control design is formulated.

II. Background

In this section, the system dynamics in input-output form is derived. A brief background on the ETC is also included.

Consider the controllable and observable uncertain nonlinear discrete time system represented as

$$x_{k+1} = f(x_k) + g(x_k)u_k , \quad (1)$$

$$y_k = Cx_k , \quad (2)$$

where $x_k \in \Re^n$, $u_k \in \Re^m$ and $y_k \in \Re^n$ represent the system state, control input and output vectors respectively while the smooth nonlinear functions $f(x_k):\Re^n \to \Re^n$, with $f(0)=0$,

Research supported in part by NSF ECCS#1128281 and Intelligent Systems Center.

$g(x_k):\Re^n \to \Re^{n\times m}$ and matrix $C\in\Re^{n\times n}$ denote the system dynamics and the output coefficient matrix respectively. The output coefficient matrix considered to be known and invertible. Here, $x=0$ being a unique equilibrium point in a compact set $\mathscr{D}\subset\Re^n$ while the system dynamics $f(x_k)$ and $g(x_k)$ are considered unknown with the following standard assumption.

Assumption 1: The unknown nonlinear function $g(\bullet)$ is uniformly bounded above in a compact set $\mathscr{D}\subset\Re^n$ such that $0<\|g(\bullet)\|\le g_M$ with $g_M>0$ is a positive known constant.

The system (1) and (2) can be transformed using the measured input/output data and represented as

$$y_{k+1}=F(y_k)+G(y_k)u_k\,, \tag{3}$$

where $F(y_k)=Cf(C^{-1}y_k)\in\Re^n$ and $G(y_k)=Cg(C^{-1}y_k)\in\Re^{n\times m}$ are the transformed system dynamics. From Assumption 1, it is routine check that $0<\|G(y_k)\|\le G_M$ with $G_M>0$ is a known positive constant. Next the background on ETC is discussed

A. Background on ETC

In a traditional output feedback based ETC frame work, the feedback information for updating the control are intermittently available based on the event-trigger condition in contrast to a traditional discrete-time frame work. Hence, the control input for an ETC system is given by

$$u_k=\begin{cases}\upsilon(y_k), \text{ Event is triggered, } k=k_i,\\ \upsilon(\breve{y}_k), \text{ Event is not triggered, } k_i\le k<k_{i+1},\end{cases} \tag{4}$$

where $\upsilon(\bullet)$ is the control input with y_k is the system output at time instant $k\in\mathbb{N}$ and $\breve{y}_k$ is the last sampled system output held by the zero-order-hold (ZOH) at the controller for $k_i\le k<k_{i+1}$ with $\{k_i\}_{i=1}^{\infty}$ is a monotonically increasing subsequence of time instants for $k\in\mathbb{N}$ at which the events are triggered. It is clear from (4) that $\breve{y}_k=y_k$ for $k=k_i$, i.e., the last sampled and held system output is updated at every trigger instant with the current sampled one.

These event-trigger instants, k_i for $i=1,2,3,\cdots$ are decided by the violation of the event-trigger condition, $e_{ET,k}\le\sigma_{ET}(y_k)$ where $e_{ET,k}$ is the event-trigger error and represented by

$$e_{ET,k}=y_k-\breve{y}_k\,, \tag{5}$$

and $\sigma_{ET}(y_k)$ is an output dependent threshold. The event-based controller (4) in the existing literature is designed without any performance criterion. Therefore, the event-based optimal control is formulated next.

B. Problem Fromulation

Consider the uncertain nonlinear discrete-time system (3). Our objective is to minimize the infinite horizon cost function,

$$V(y_k)=\sum_{j=k}^{\infty}r(y_j,u_j)\,, \tag{6}$$

by designing an optimal feedback control policy in an event-triggered control framework, where $r(y_k,u_k)$ is the cost-to-go function at each time k and defined as $r(y_k,u_k)=Q(y_k)+u_k^TRu_k$ with $Q(y_k)\in\Re$ is a positive definite function to penalize the system output, y_k, and R is a positive definite matrix to penalize the control input, u_k. An initial admissible control input is needed to keep the cost function finite.

The cost function (6) can be written as

$$V(y_k)=r(y_k,u_k)+\sum_{j=k+1}^{\infty}r(y_j,u_j)=Q(y_k)+u_k^TRu_k+V(y_{k+1}), \tag{7}$$

with $V(y_{k+1})$ is the value function from time instant $k+1$ onwards. According to Bellman's principle of optimality [8], the optimal cost, $V^*(y_k)$, satisfies the discrete-time Hamilton Jacobi Bellman (HJB) equation

$$V^*(y_k)=\min_{u_k}\left\{Q(y_k)+u_k^TRu_k+V^*(y_{k+1})\right\}, \tag{8}$$

where $V^*(y_k)$ is the optimal cost at time instant, k, and $V^*(y_{k+1})$ is the optimal cost for $k+1$ onwards. The optimal control policy u_k^* is derived from the stationarity condition [7], $\partial V^*(y_k)/\partial u_k=0$, and it is found to be

$$u_k^*=-(1/2)R^{-1}G^T(y_k)\partial V^*(y_{k+1})/\partial y_{k+1}\,. \tag{9}$$

Substituting (9) in (8), the discrete-time HJB equation is obtained.

In general, an exact analytical solution to the HJB equation (8) is not possible. Further, it is evident from (9) that the optimal control input requires control coefficient function $G(y_k)$ and the system output y_{k+1} at the time instant, k. This precludes the computation of the control input (9) at the current time instant, k, in a forward-in-time manner. Further, in the event-based optimal control, the output vector is sent intermittently to the controller at event-triggered instants thus complicating the approximation process. Next, the function approximation by NN in the context of event-based sampling is revisited.

C. NN Approximation with Event-based Sampling

The following theorem extends the approximation property of NNs for event-based sampling.

Theorem 1: Let $h(y_k):\Re^n\to\Re^n$, be smooth and continuous function in a compact set $\mathscr{D}\subset\Re^n$. Then, there exists a NN with a sufficient number of neurons such that the function $h(y_k)$ can be approximated with constant weights and event-based activation function defined as

$$h(y_k)=W^T\sigma(\breve{y}_k)+\varepsilon_e(\breve{y}_k)\,, \tag{10}$$

where $W\in\Re^{l\times n}$ is the target NN weight matrix with l being the number of hidden-layer neurons, $\sigma(\breve{y}_k)\in\Re^l$ is bounded event-based activation function, and $\varepsilon_e(\breve{y}_k)\in\Re^n$ being the event-driven NN reconstruction error with $\breve{y}_k\in\Re^n$ is the latest available event-sampled variable at the time instant, k_i.

Remark 1: From Theorem 1, a small reconstruction error $\varepsilon_e(\breve{y}_k)$ can be achieved by increasing the frequency of triggered events. Hence, a compromise has to be reached between NN reconstruction error and computation in an event-triggered control scheme.

III. EVENT-BASED OPTIMAL CONTROL DESIGN

In this section, we propose the near optimal controller design in the event-triggered framework.

A. *Proposed Solution*

The optimal event-triggered control system is shown in the Figure 1. An actor-critic framework is proposed in the event-triggered control paradigm. The value function which is the solution of the HJB equation is approximated by the critic NN and the control input by the actor NN which minimizes the HJB equation. Further, the control coefficient function $G(y_k)$ in (9) is relaxed by using a NN based identifier.

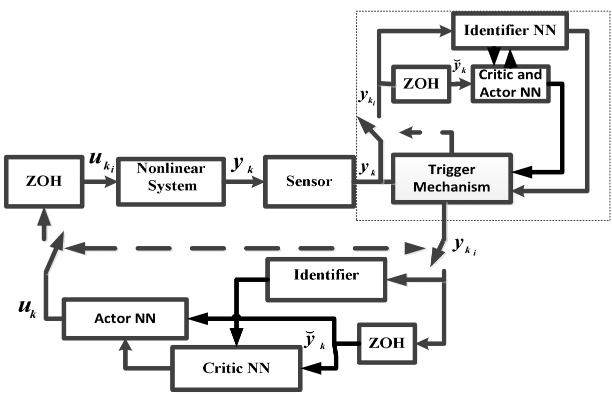

Fig. 1. Block diagram of an event-triggered near optimal control system.

All the three NNs use event-sampled states as inputs and the sampling instants are decided by the trigger mechanism. The event-trigger mechanism consists of a mirror image of the actor-critic NN, which operates in synchronism with the controller NNs, to evaluate the event-trigger condition at every sampling instant. The system output is transmitted at the violation of the event-trigger condition defined later. Further, the NN identifier, the last held output at the ZOH, the control input and NN weights are updated at every trigger instant for $k = k_i$. Similar to the case of a traditional ETC the event-trigger error is reset to zero, i.e.,

$$e_{ET,k} = 0 \text{ for } k = k_i, \tag{11}$$

for the next iteration. In the next subsections, the identifier and controller designs are presented.

B. *Identifier Design*

By using the approximation property of the dynamic feed forward NN [6], the nonlinear system in (3) can be represented as

$$y_{k+1} = W_F^T \sigma_F(y_k) + W_G^T \sigma_G(y_k) u_k + \varepsilon_F + \varepsilon_G u_k = W_I^T \sigma_I(y_k)\bar{u}_k + \varepsilon_{I,k}, \tag{12}$$

where $W_I = \begin{bmatrix} W_F^T & W_G^T \end{bmatrix}^T \in \Re^{(m+1)l_I \times n}$ denotes the target NN weights of system identifier, $\sigma_I(y_k) = diag\{\sigma_F(y_k), \sigma_G(y_k)\} \in \Re^{(m+1)l_I \times (m+1)}$ represents the NN activation function with l_I is the number of neurons, $\varepsilon_I = \varepsilon_F + \varepsilon_G u_k$ is the reconstruction error and $\bar{u}_k = \begin{bmatrix} 1 & u_k^T \end{bmatrix}^T \in \Re^{m+1}$ is the augmented control input. Then, the following standard assumption for the NN holds.

Assumption 2[6]: The target weight vector, W_I, the activation function, $\sigma_I(y_k)$, and the reconstruction error, ε_I, of the NN are bounded in a compact set $\mathscr{D} \subset \Re^n$ such that $\|W_I\| \le W_{I,M}$, $\|\sigma_I(y_k)\| \le \sigma_{I,M}$ and $\|\varepsilon_I\| \le \varepsilon_M$ with W_M, σ_M and ε_M are positive constants .

Next, defining the NN identifier dynamics as

$$\hat{y}_{k+1} = \hat{F}(\breve{y}_k) + \hat{G}(\breve{y}_k)u_k = \hat{W}_{I,k}^T \sigma_I(\breve{y}_k)\bar{u}_k, \tag{13}$$

where $\hat{y}_k \in \Re^n$ is the identifier state vector , $\hat{F}(\breve{y}_k) \in \Re^n$ and $\hat{G}(\breve{y}_k) \in \Re^{n \times m}$ are the identifier dynamics at the time instant k and a function of last held system output , $\breve{y}$. Further, $\hat{W}_{I,k} = \begin{bmatrix} \hat{W}_{F,k}^T & \hat{W}_{G,k}^T \end{bmatrix}^T \Re^{(m+1)l_I \times n}$ denotes the estimated weights of identifier, $\sigma_I(\breve{y}_k) = diag\{\sigma_F(\breve{y}_k), \sigma_G(\breve{y}_k)\} \in \Re^{(m+1)l_I \times (m+1)}$ is the event-based activation function. As per the proposed scheme $\hat{y}_k = y_k$ and $\breve{y}_k = y_k$ for $k = k_i$, and, hence, the dynamics (13) are represented by

$$\hat{y}_{k+1} = \begin{cases} \hat{W}_{I,k}^T \sigma_I(y_k)\bar{u}_k, \text{ event is triggered, } k = k_i, \\ \hat{W}_{I,k}^T \sigma_I(\breve{y}_k)\bar{u}_k, \text{event is not triggered, } k_i \le k < k_{i+1}. \end{cases} \tag{14}$$

Defining the identification error as $e_{I,k} = y_k - \hat{y}_k$, the identification error dynamics become

$$e_{I,k+1} = \tilde{W}_{I,k}^T \sigma_I(y_k)\bar{u}_k + \hat{W}_{I,k}^T\left(\sigma_I(y_k) - \sigma_I(\breve{y}_k)\right)\bar{u}_k + \varepsilon_{I,k}, \tag{15}$$

with $\tilde{W}_{I,k} = W_I - \hat{W}_{I,k}$ is the identifier NN weight estimation error. Further, the dynamics of $e_{I,k}$ for $k = k_i$ from (15) become

$$e_{I,k+1} = y_{k+1} - \hat{y}_{k+1} = \tilde{W}_{I,k}^T \sigma_I(y_k)\bar{u}_k + \varepsilon_{I,k}, \text{ for } k = k_i. \tag{16}$$

Next, the aperiodic update law is selected in the context of event-based sampling to save computational load as

$$\hat{W}_{I,k+1} = \begin{cases} \hat{W}_{I,k} + \alpha_I \dfrac{\sigma_I(y_k)\bar{u}_k e_{I,k+1}^T}{[\sigma_I(y_k)\bar{u}_k]^T[\sigma_I(y_k)\bar{u}_k] + 1}, \text{for } k = k_i, \\ \hat{W}_{I,k}, \qquad \text{for } k_i \le k < k_{i+1}, \end{cases} \tag{17}$$

with α_I is the learning gain and $e_{I,k+1}$ is defined as in (16).

C. *Controller Design*

In this section, the near optimal controller design in the context of event-based sampling is given via actor-critic framework. First, the critic NN design is presented.

1) *Critic Design*

In this subsection, we learn the solution of the HJB equation function by using the critic NN. Consider the ideal HJB equation (8) rewritten as

$$0 = V^*(y_{k+1}) + Q(y_k) + R(u_k^*) - V^*(y_k)\,. \tag{18}$$

By using the dynamic feed forward NN [6], the optimal value function can be written as

$$V^*(y_k) = W_V^T \varphi(y_k) + \varepsilon_{V,k}\,, \tag{19}$$

where $W_V \in \Re^{l_V}$ is the unknown constant target critic NN weights, $\varphi(y_k) \in \Re^{l_V}$ is the activation functions, $\varepsilon_{V,k} \in \Re$ is the NN reconstruction errors with l_V is the number of neurons in the network. Similar to Assumption 2 the target NN weights, activation functions and the reconstruction errors of the critic NN are assumed to be upper bounded by $\|W_V\| \le W_{V,\mathrm{M}}$, $\|\varphi(y_k)\| \le \varphi_{\mathrm{M}}$, and $|\varepsilon_{V,k}| \le \varepsilon_{V,\mathrm{M}}$ with $W_{V,\mathrm{M}}$, φ_{M} and $\varepsilon_{V,\mathrm{M}}$ are positive constants. In addition, in this work, the gradient of the reconstruction error is also considered to be bounded above by $\|\partial\varepsilon_{V,k}/\partial y_{k+1}\| \le \varepsilon'_{V,\mathrm{M}}$, with $\varepsilon'_{V,\mathrm{M}}$ a positive constant [9]. Then, substituting (19) into (18), (18) can be represented as

$$0 = W_V^T \Delta\varphi(y_k) + Q(y_k) + u_k^{*T} R u_k^* + \Delta\varepsilon_{V,k}\,, \tag{20}$$

where $\Delta\varepsilon_{V,k} = \varepsilon_{V,k+1} - \varepsilon_{V,k}$ and $\Delta\varphi(y_k) = \varphi(y_{k+1}) - \varphi(y_k)$.

Now, the estimated value function by the critic NN with the event-based availability of the system output can be represented as

$$\hat{V}(y_k) = \hat{W}_{V,k}^T \varphi(\breve{y}_k)\,, \tag{21}$$

where $\hat{W}_{V,k}^T \in \Re^{l_V}$ is the estimate of the target NN weight of the value function, $\varphi(\breve{y}_k) \in \Re^{l_V}$ is the event-based activation function. The event-based activation function is selected such that $\|\sigma(0)\| = 0$ for $\|y_k\| = 0$ to ensure $\hat{V}(0) = 0$.

Because of the NN approximation (21) and the event-based availability of the system state, $\breve{y}_k$ for $k_i \le k < k_{i+1}$, the ideal HJB equation (20) does not hold. Therefore the temporal difference (TD) error associated with the estimated HJB equation can be written as

$$e_{TD,k} = \hat{W}_{V,k}^T \Delta\varphi(\breve{y}_k) + Q(\breve{y}_k) + u_k^T R u_k\,. \tag{22}$$

At the trigger instant, $k = k_i$, the temporal difference (TD) error can be written as

$$e_{TD,k} = \hat{W}_{V,k}^T \Delta\varphi(y_k) + Q(y_k) + u_k^T R u_k, \quad \text{for } k = k_i\,, \tag{23}$$

where $\Delta\varphi(\breve{y}_k) = \varphi(\breve{y}_{k+1}) - \varphi(\breve{y}_k)$.

Now, selecting the aperiodic update law of critic NN as

$$\hat{W}_{V,k+1} = \begin{cases} \hat{W}_{V,k} - \alpha_V \dfrac{\Delta\varphi(y_k) e_{TD,k}}{\Delta\varphi^T(y_k)\Delta\varphi(y_k)+1}, & \text{for } k = k_i, \\ \hat{W}_{V,k}, & \text{for } k_i \le k < k_{i+1}, \end{cases} \tag{24}$$

with $\alpha_V > 0$ is the learning gain .

Remark 2: The update law (24) , especially the TD error, $e_{TD,k}$, uses y_{k+1} to update the weight at $k+1$. Hence, the system output y_{k+1} is also sent to the controller at $k+1$ once an event is trigged at the time instant k. After receiving the output at $k+1$, the critic NN weight for $k+1$ is updated. Next, the design of the actor NN is presented .

B. Actor NN Design

In this subsection, we approximate the optimal control policy through the actor NN.

Consider the optimal control input (9). By using the dynamic the feed forward NN [6] (9) can be written as

$$u_k^* = W_u^T \sigma_u(y_k) + \varepsilon_{u,k}\,, \tag{25}$$

where $W_u \in \Re^{l_u \times m}$ is the target action NN weights, $\sigma_u(y_k) \in \Re^{l_u}$ denotes the activation function of the action NN and $\varepsilon_{u,k} \in \Re^m$ is the reconstruction error with l_u is the number of neurons in the network. As per Assumption 2 the target action NN weights, activation function, reconstruction error satisfies $\|W_u\| \le W_{u,M}$, $\|\sigma_u(y_k)\| \le \sigma_{u,M}$ and $\|\varepsilon_{u,k}\| \le \varepsilon_{u,M}$ respectively.

Assumption 3: The NN activation function is Lipschitz continuous in a compact set $\mathcal{D} \subset \Re^n$. This means for every $\mathcal{D}$ there exists a Lipschitz constant C_{σ_u} such that $\|\sigma_u(y_k) - \sigma_u(\hat{y}_k)\| \le C_{\sigma_u}\|y_k - \hat{y}_k\|$.

Moreover, the ideal optimal control input (9) by using (19), by forwarding one step, becomes

$$u_k^* = -(1/2)R^{-1}G^T(y_k)\nabla\varphi^T(y_{k+1})W_V - (1/2)R^{-1}G^T(y_k)\nabla\varepsilon_{V,k+1}\,. \tag{26}$$

It is clear that both the optimal control inputs (25) and (26) should be equal.

Next, the estimated optimal control input in the event-trigger context can be represented by using the actor NN weights as

$$u_k = \hat{W}_{u,k}^T \sigma_u(\breve{y}_k)\,, \text{ for all } k \in \mathbb{N}\,, \tag{27}$$

where $\hat{W}_{u,k} \in \Re^{l_u \times m}$ is the estimated actor NN weights and $\sigma_u(\breve{y}_k)$ is the event-based activation function. Alternatively, using the gradient of the estimated value function (21), the estimated control input, $u_{V,k}$, can also be written as

$$u_{v,k} = -(1/2)R^{-1}\hat{G}^T(\breve{y}_k)\nabla\varphi^T(\breve{y}_{k+1})\hat{W}_{V,k}\,, \text{ for all } k \in \mathbb{N}\,, \tag{28}$$

where ∇ denotes the gradient, $\hat{G}(\hat{y})$ is the estimated control coefficient matrix from the NN-based identifier.

It is clear that the control policy, u_k, (27) and the control policy, $u_{v,k}$, (28) are not equal and, hence, the control input estimation error $e_{u,k}$, which is the difference between (27) and (28), is represented as

$$e_{u,k} = \hat{W}_{u,k}^T \sigma_u(\breve{y}_k) + (1/2)R^{-1}\hat{G}^T(\breve{y}_k)\nabla\varphi^T(\breve{y}_{k+1})\hat{W}_{V,k}\,. \tag{29}$$

Then from (29), $e_{u,k}$ for $k = k_i$ can be written as

$$e_{u,k} = \hat{W}_{u,k}^T \sigma_u(y_k) + (1/2)R^{-1}\hat{G}^T(y_k)\nabla\varphi^T(y_{k+1})\hat{W}_{V,k}\,. \tag{30}$$

Similar to (24), the update law of action NN is selected as

$$\hat{W}_{u,k+1} = \begin{cases} \hat{W}_{u,k} - \alpha_u \dfrac{\sigma_u(y_k) e_{u,k}^T}{\sigma_u^T(y_k)\sigma_u(y_k)+1}, & \text{for } k = k_i \\ \hat{W}_{u,k}, & \text{for } k_i \le k < k_{i+1} \end{cases}, \tag{31}$$

with α_u is the learning gain and $e_{u,k}$ is as defined in (30).

Next, the flowchart given in Fig. 2 explains the implementation of the near optimal event-triggered control design presented earlier.

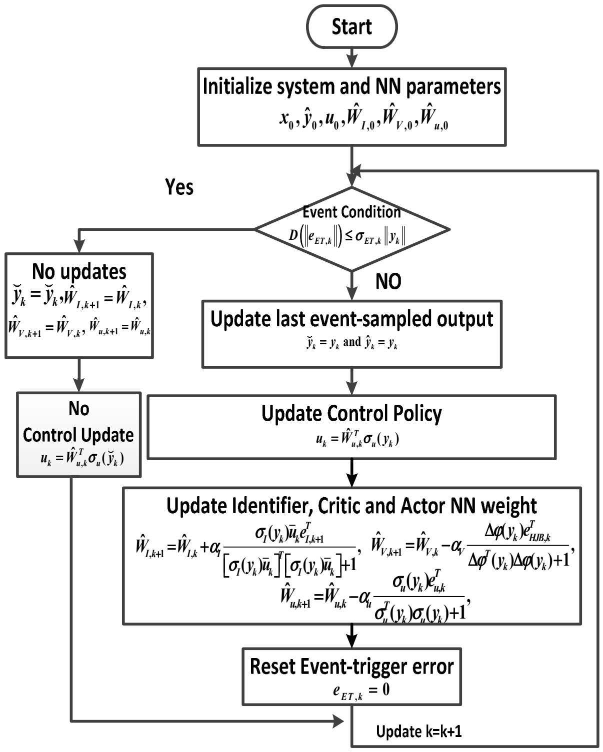

Fig. 2. Flowchart for the proposed event-triggered system

To start with, the system states, identifier states and the NN weights are initialized. Now, in the second step, the event-trigger condition is evaluated using the following condition

$$D\left(\left\|e_{ET,k}\right\|\right) \le \sigma_{ET,k} \left\|y_k\right\|, \tag{32}$$

where $\sigma_{ET,k} = \sqrt{(1-2K^2)\Gamma_{ET} \Big/ 4G_M^2 C_{\sigma_u}^2 \left\|\hat{W}_{u,k}\right\|^2}$, (33)

with $0 < \Gamma_{ET} < 1$, G_M is the upper bound for the control coefficient matrix function $G(y_k)$, C_{σ_u} is the Lipschitz constant for the actor NN activation function and $0 < K < \sqrt{1/2}$ is a constant satisfies the inequality $\left\|F(y_k)+G(y_k)u_k^*\right\|^2 \le K\left\|y_k\right\|^2$. The dead-zone operator $D(\bullet)$ is defined as

$$D\left(\left\|e_{ET,k}\right\|\right) = \begin{cases} \left\|e_{ET,k}\right\|, & \left\|y_k\right\| > B_M^y, \\ 0, & \text{otherwise}, \end{cases} \tag{34}$$

where B_M^y is the ultimate bound for the system output. If the condition is satisfied there will be no update of the control input and the NN weights. The system runs with the previously held updated control input.

On the other hand, once the event-trigger condition is violated, the last held system outputs and the control inputs are updated and fed back to the system. To update the NN weights in the next step both at the controller and trigger mechanism, the system output y_{k+1} is made available at the controller. Finally, the event-trigger error is reset to zero and time is incremented. Next the main result is presented.

Theorem 2: Consider the nonlinear discrete-time system (3), the NN identifier (14), critic NN (21) and actor NN (27). Let u_0 be an initial stabilizing control policy for (3) and let the Assumptions 1 through 3 hold and the identifier, critic and actor NN weights, $\hat{W}_{I,0}$, $\hat{W}_{V,0}$ and $\hat{W}_{u,0}$ be initialized in $\mathscr{D} \subset \Re^n$. Then, there exist positive constants α_I , α_V and α_u such that the closed-loop event-triggered system output vector, y_k , for any initial condition in the compact set $\mathscr{D} \subset \Re^n$, the identifier, critic and actor NN weight estimation errors respectively, $\tilde{W}_{I,k}$, $\tilde{W}_{V,k}$ and $\tilde{W}_{u,k}$, are uniformly ultimate bounded (UUB) provided the system outputs are sent to the controller and all the NN weights are updated using (17), (24) and (31) respectively through the violation of the event-trigger condition (32). Further, $\left\|V^* - \hat{V}\right\| \le b_V$ and actual control input is bounded close to optimal control input such that $\left\|u^* - u\right\| \le b_u$ with $b_V = b_{\tilde{W}_V}\varphi_M + 2W_{V,M}\varphi_M + \varepsilon_{v,M}$ and $b_u = 3b_{\tilde{W}_u}\sigma_{u,M} + 2W_{u,M}\sigma_{u,M} + \varepsilon_{u,M}$ are small positive constants with $b_{\tilde{W}_V}$ and $b_{\tilde{W}_u}$ respectively are the ultimate bounds for $\tilde{W}_{V,k}$ and $\tilde{W}_{u,k}$ when the design parameters are selected as

$$0 < \alpha_u \le \frac{\alpha_V \sigma_{u,\mathrm{m}}^2 (\sigma_{u,M}^2 + 1)\Delta\varphi_{\mathrm{m}}^2 \Gamma_V}{8(2W_{I,M}^2 + 1)\sigma_{u,M}^2 \nabla\varphi_M^2 \sigma_{I,M}^2 \lambda_{\max}^2 (R^{-1})(\Delta\varphi_M^2 + 1)} < 1 ,$$

$0 < \alpha_I < \dfrac{1}{2}$, and $0 < \alpha_V < 3\left(\Delta\bar{\varphi}_m^2 + 1\right) \big/ 2\Delta\bar{\varphi}_m^2$.

Remark 3: The event-trigger condition (32) is designed using the Lyapunov direct method such that all the closed-loop system parameters are UUB which is stated in the Theorem 2. Further, the dead zone operator (34) used with the event-trigger condition stops the unnecessary triggering once the system output is in the ultimate bound B_M^y .

Remark 4: It is important to notice from (2) that the system output is a constant multiplier of the system state. Hence, the system state also remains UUB by Theorem 2.

IV. Simulation Results

In this section, a practical two-link robot has been considered for simulation. The dynamics of the two-link robot is given as $x_{k+1} = f(x_k) + g(x_k)u_k$, $y_k = Cx_k$ with internal dynamics $f(x_k)$ and control coefficient matrix $g(x_k)$ are given as in [11]. The output coefficient matrix is selected as $C = [0\ 1\ 0\ 0.2;\ 1\ 0\ 0\ 0.2;\ 0\ 1\ 0.1\ 0;\ 0\ 0\ 1\ 0.1]$. The value function is selected in quadratic form as $Q(y_k) = y_k^T Q_y y_k$ with $Q_y = 0.001 * I_{4\times4}$ and $R = 0.001 * I_{2\times2}$ with I is the identity matrix. Three dynamic feed-forward NNs are used with online learning and no offline training is required. The activation function for identifier and the actor NN are chosen as $\tanh(\bullet)$ and for the critic network $\varphi(y)=[y_{1,k}^2; y_{2,k}^2; y_{3,k}^2; y_{4,k}^2; y_{1,k}y_{2,k} \cdots; y_{1,k}^4; y_{2,k}^4; \cdots y_{1,k}^3 y_{2,k}; \cdots y_{1,k}^2 y_{2,k} y_{3,k}; \cdots; y_{1,k} y_{2,k} y_{3,k} y_{4,k}]$. The number of neurons for the identifier is 79, and for critic and actor NN are 39 each. The learning rates for the NN tuning are selected as $\alpha_I = 0.095$, $\alpha_V = 0.05$, $\alpha_u = 0.005$ and $\Gamma_{ET} = 0.99$. The ultimate bound selected for the system output is 0.001.

The performances of the event-triggered optimal control are shown in Fig. 3 to Fig. 5. Convergence of the system state and control input are shown in Fig. 3 (a) and (b) respectively. The cumulative number of triggered events for a simulation interval of 30 sec with a sampling of 0.001 sec is shown in Fig. 3(c). From the plot, the number of events with the proposed scheme is around 3501 (out of 30000 sampling instants). The total number of additions/subtractions and multiplications for the traditional discrete-time is 930,000 whereas for the proposed scheme is 252,072. This implies a reduction in computational load of the event-triggered system of 73% when compared to a traditional discrete time system. The plot in Fig. 3 (d) show inter- event time.

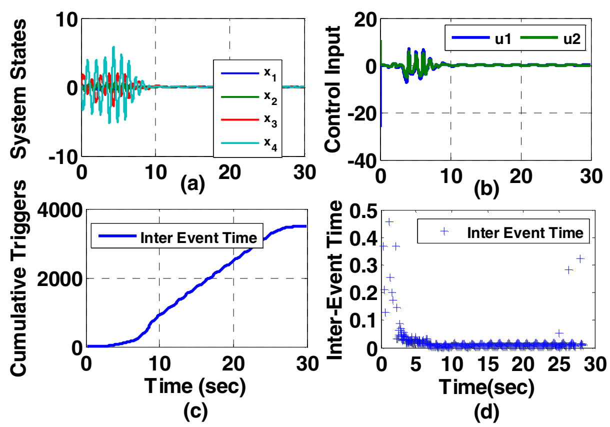

Fig. 3. Convergence of (a) system states, and (b) optimal control input. (c) cumulative number of triggered events, and (d) inter-event time.

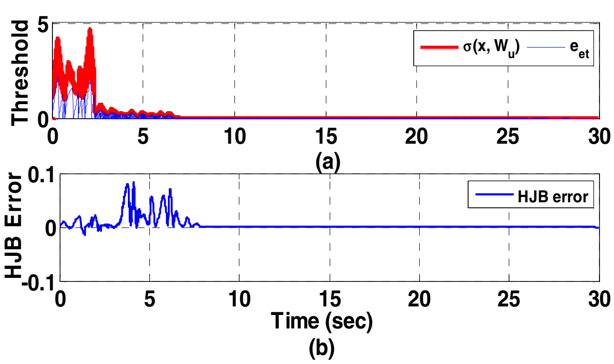

Fig. 4. (a) Triggering threshold with event-trigger error; (b) Temporal difference error.

Fig. 4 (a) shows the evolution of the event threshold (red) over time along with the event-trigger error (blue). From Fig. 4 (a) it is evident that once the event-trigger error reaches the threshold an event is triggered and the error reset to zero. This ensures the stability of the closed-loop system. Further, the convergence of the TD error is plotted in Figure 4(b) confirming the optimality is achieved. The boundedness of actual NN weights is shown in Fig. 5.

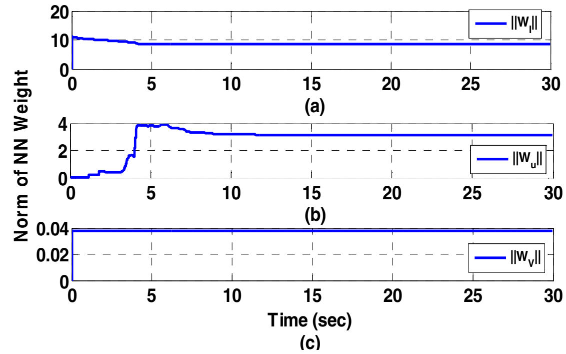

Fig. 5. Convergence of NN weights.

V. Conclusion

In this paper, a near optimal event-triggered control of an uncertain nonlinear discrete-time system in affine form is introduced. An actor-critic frame work is presented to solve the optimal control problem with event-based availability of the system output at the controller. The novel event-trigger condition generated required number of events at the initial learning phase to achieve a small error in approximation and subsequently saved the computation by fewer updates in the control law. The near optimality is achieved with complete unknown system dynamics. The effectiveness of the controller is validated by using simulation studies.

References

[1] K. J. Åström and B. M. Bernhardsson, "Comparison of periodic and event based sampling for the first order stochastic systems," in *Proc. IFAC World Conf.*, 1999, pp. 301–306.

[2] P. Tabuada, "Event-triggered real-time scheduling of stabilizing control tasks," *IEEE Trans. on Autom. Control*, vol. 52, pp. 1680-1685, 2007.

[3] M. C. F. Donkers and W. P. M. H. Heemels, "Output-based event-triggered control with guaranteed L•-gain and improved and decentralized event triggering," *IEEE Trans. on Autom. Control*, vol. 57, pp. 1362–1376, 2012.

[4] A. Molin and S. Hirche, "On the optimality of cerntainty equivalence for event-triggered control systems," *IEEE Trans. Automat. Contr.*, vol. 58, pp. 470-474, 2013.

[5] A. Sahoo, H. Xu, and S. Jagannathan, "Neural network based adaptive event-triggered control of affine nonlinear discrete-time systems with unknown internal dynamics," in *proc. Amer. Contr. Conf.*, 2013, pp. 6433-6438.

[6] S. Jagannathan, '*Neural network of nonlinear discrete-time system,"* CRC Press, 2006.

[7] F. L. Lewis and V. L. Syrmos, *Optimal Control*, 2nd edition. New York: Wiley, 1995.

[8] D. P. Bertsekas, *Dynamic Programming and Optimal Control*, 2nd ed. Belmonth, MA: Athena Scientific, 2000.

[9] T. Dierks and S. Jagannathan, "Online optimal control of affine nonlinear discrete-time system with unknown internal dynamics by using time-based policy update," *IEEE Trans. Neur. Net. And Learn. Syst.*, vol. 23, pp. 1118-1129, 2012.

[10] R. Beard, "Improving the closed-loop performance of nonlinear systems," Ph.D. dissertation, Rensselaer Poly Institute, USA, 1995.

[11] Z. Chen and S. Jagannathan, "Generalized Hamilton-Jacobi-Bellman formulation-based neural netwrok control of affine nonlinear discrete-time system," *IEEE Transac. Neural Networks,* vol. 19 pp. 90-106, 2008.