Summer 2015

Event sampled optimal adaptive regulation of linear and a class of nonlinear systems

Avimanyu Sahoo

Department: Electrical and Computer Engineering

EVENT SAMPLED OPTIMAL ADAPTIVE REGULATION OF LINEAR AND A CLASS OF NONLINEAR SYSTEMS

by

AVIMANYU SAHOO

A DISSERTATION

Presented to the Faculty of the Graduate School of the

MISSOURI UNIVERSITY OF SCIENCE AND TECHNOLOGY

In Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

in

ELECTRICAL ENGINEERING

2015

Approved by

Jagannathan Sarangapani, Advisor
S. N. Balakrishnan
Sanjay Kumar Madria
Maciej Zawodniok
Sriram Chellappan

PUBLICATION DISSERTATION OPTION

This dissertation consists of the following five articles that have been submitted for publication. The papers are formatted according to Missouri University of Science and Technology specifications.

Pages 17-66 are under review with AUTOMATICA.

Pages 67-116 are conditionally accepted to IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS.

Pages 117-167 are conditionally accepted in IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS.

Pages 168-215 Accepted in IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS.

Pages 216-262 are submitted to IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS.

This work was supported in part by the National Science Foundation grants ECCS #1406533 and in part by Intelligent Systems Center, Missouri S&T, Rolla, MO, USA.

ABSTRACT

In networked control systems (NCS), wherein a communication network is used to close the feedback loop, the transmission of feedback signals and execution of the controller is currently carried out at periodic sampling instants. Thus, this scheme requires a significant computational power and network bandwidth. In contrast, the event-based aperiodic sampling and control, which is introduced recently, appears to relieve the computational burden and high network resource utilization. Therefore, in this dissertation, a suite of novel event sampled adaptive regulation schemes in both discrete and continuous time domain for uncertain linear and nonlinear systems are designed.

Event sampled Q-learning and adaptive/neuro dynamic programming (ADP) schemes without value and policy iterations are utilized for the linear and nonlinear systems, respectively, in both the time domains. Neural networks (NN) are employed as approximators for nonlinear systems and, hence, the universal approximation property of NN in the event-sampled framework is introduced. The tuning of the parameters and the NN weights are carried out in an aperiodic manner at the event sampled instants leading to a further saving in computation when compared to traditional NN based control.

The adaptive regulator when applied on a linear NCS with time-varying network delays and packet losses shows a 30% and 56% reduction in computation and network bandwidth usage, respectively. In case of nonlinear NCS with event sampled ADP based regulator, a reduction of 27% and 66% is observed when compared to periodic sampled schemes. The sampling and transmission instants are determined through adaptive event sampling conditions derived using Lyapunov technique by viewing the closed-loop event sampled linear and nonlinear systems as switched and/or impulsive dynamical systems.

ACKNOWLEDGEMENTS

I would like to express my special appreciation and sincere gratitude to my advisor, Professor Jagannathan Saragapani, for encouraging my research and allowing me to grow as a researcher. I would like to thank him for his effective and timely supervision, help and support during the entire research duration. Without his patient and constant guidance this dissertation would not have been possible. His advice on both research as well as on my career have been invaluable. I would also like to thank my committee members Dr. S. N. Ballakrishnan, Dr. Sanjay Kumar Madria, Dr. Maciej Zawodniok, and Dr. Sriram Chellappan for serving as my committee member. I would like to thank all of them for their valuable comments and suggestions.

I would like to thank Dr. Hao Xu, for his timely help, valuable suggestions and discussions concerning to my dissertation work. In addition, I would like to thank the National Science Foundation (NSF) and Intelligent System Center (ISC) for providing financial support throughout my Ph.D. study. I would like to thank all faculty and staff of Missouri University of Science and Technology for their direct and/or indirect help in making this dissertation successful. I would like to thank my colleagues, Dr. Hassan Zargarzadeh, Dr. Qiming Zhao, as well as many other friends at the Embedded Systems and Networking Laboratory, who made my Ph.D. life more fun and interesting.

Finally, I want to express my special thanks to my family. I am grateful to my mother and father for all of their sacrifices and prayer which made me to sustain this far. I would like to thank my brothers and other family members for their love, care and encouragement, without which I would have not been at this stage of my life.

TABLE OF CONTENTS

SECTION

APPENDICES

LIST OF ILLUSTRATIONS

PAPER III

PAPER IV

PAPER V

LIST OF TABLES

SECTION

1. INTRODUCTION

The advent of embedded processors spurred research on digital implementation of the controllers. Traditionally, the sampled data [1] and discrete-time control [2] frame works are used for this purpose because of the well-developed theory. In the sample data system approach [1], a continuous-time plant is controlled by a discrete time controller whereas in a discrete time control [2] the system itself operates under discrete-time mode. In both the schemes, a periodic, fixed sampling time, decided a priori, is used for sampling the feedback signals and controller execution. This fixed sampling time is, in general, governed by the well-known Nyquist sampling criterion by considering the worst case scenario. The wide spread application of this sampling scheme is due to the ease of analysis with the numerous available techniques in the literature.

On the other hand, this periodic sampling leads to ineffective resource utilization [3] with higher control cost for dynamic systems having limited computational capability. The problem aggravates in the case of systems with shared digital communication network in the feedback loop, referred to as networked control systems (NCS) [4]-[7], due to limited bandwidth. The periodic sampling and transmission further escalates the problem with network congestion leading to longer network induced delays. Furthermore, this periodic sampling and transmission of feedback data and controller execution is redundant in situations when there is no significant change in overall system performance and the system is operating with desired output.

As an alternative, to alleviate the burden of needless computational load and network congestion, various sampling schemes [8]-[12] were proposed. In the recent

times, performance based sampling schemes are developed to reduce the computational cost and formally referred to as "event-triggered control" [13]-[25]. This sampling scheme decides the transmission and controller execution instants when there is a significant change in the system state or output errors that can either jeopardize the stability or deteriorate the desired performance. This requires an additional hardware device, referred to as trigger mechanism, to evaluate the event-triggering condition which orchestrates the sampling instants or simply events. Since, the objective of this sampling is controller execution, not for signal reconstruction, this is equally applicable for both continuous [14]-[23] and discrete time systems [24]-[26] to either regulate the system state vector to zero [14]-[17] or track a desired trajectory [18]. A general layout of a discrete time event-triggered system is shown on Figure 1.1.

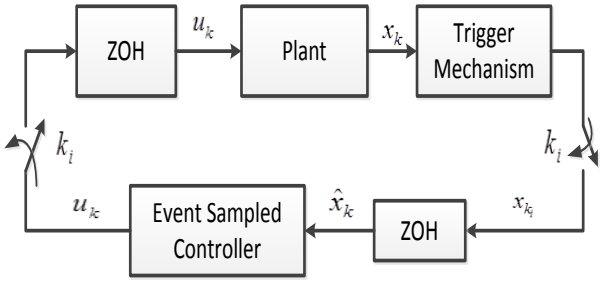

Figure 1.1 Block diagram of the discrete time event sampled control system.

In the case of a continuous-time system, the sensor measures the system state or output vectors continuously and the trigger mechanism determine the sampling instants by evaluating the event-triggering condition [14]-[18]. The event-triggering condition is usually a function of the state error, referred to as event-triggering error, and a suitably designed state dependent threshold [14]. The feedback signals are transmitted and control

is executed when the event-triggering error exceeds the threshold. Lyapunov stability technique or its extensions are used as the work horse to design the event-triggering condition that ensures the stability and desired performance of the system.

In the discrete-time case [24]-[26], the sensor measures the system state or output and the trigger mechanism evaluates the event-triggering condition at every periodic sampling instant and a decision is made whether to transmit or not. The feedback data are transmitted and controller is executed only at the violation of the event-triggering condition. In both continuous and discrete time cases, the event-triggering instants or simply the events turn out to be aperiodic in nature and, hence, save computational load and bandwidth usage. These inherent advantages of event-triggered control is proven to be more beneficial in large scale systems such a decentralized systems [28]-[31], multi agent systems [32]-[33], and cyber-physical systems [34] to name a few.

A similar approach called self-triggered control [35]-[36] is also developed for systems where the extra hardware for the trigger mechanism is hard to implement. This software based scheme, which is a special case of the event-triggered control, predicts the sampling instants by using the previously sampled data and the dynamics of the system. Hence, a continuous evaluation of the event-triggering condition is not necessary. The analysis of the self-triggered [35]-[36] system is similar to that of the event-triggered control and is outside the scope of this dissertation. Next, a detailed overview of the event-triggered control schemes available in the literature is presented and the motivation behind this research is discussed.

1.1 OVERVIEW OF THE EVENT-TRIGGERED CONTROL

The study of aperiodic sampling for *sampled data control* dates back to the late fifties and early sixties [8]-[12] and was first studied in [8] for quantized systems to share the communication channel without increasing its bandwidth. Moreover, a state based adaptive sampling method for sampled data servo mechanism is proposed in [9] where the adaptive sampling rate is controlled by the absolute value of the first derivative of the error signal. Lately, this aperiodic state dependent sampling is studied under various names, such as, multi rate sampling [10], interrupt driven triggering [11], level triggered sampling [12]. Recently, this scheme is studied under a formal name of *event-triggered sampling* [14]-[41], [43]-[44] and various theoretical [3], [14] and experimental [13], [16] results emphasizing its inherent advantages, in computation and communication saving, are available in the literature.

In the last few years, theoretical results started to appear in the literature for both deterministic [14]-[35] and stochastic [38] event-triggered control and thereafter various controller designs are introduced. A majority of the theoretical results on event-triggered control both for linear and nonlinear systems are available for deterministic systems [14]-[35]. In general, the emulation-based approach [14]-[15], [18], [30] is used for the event-triggered system design. Emulation based design in the sense that the continuous controller is presumed to be stabilizing and an event-triggering condition is developed to implement the controller such that the stability and certain level of performance are maintained. In the earlier works [14]-[15] the system was assumed to be input to state stable (ISS) [49] with respect to the measurement error, and event-triggering conditions are designed to reduce computation and guarantee asymptotic stability. A non-zero

positive lower bound on the inter-event times is also guaranteed to avoid accumulation point and Zeno behavior. On the other hand, the stringent ISS assumption is relaxed by assuming the asymptotic stability of the continuous system by the authors in [18], [20]. The event-triggered control approach is also extended to accommodate other design considerations, such as, output feedback design [20], [23], [39], decentralized designs [28]-[31], and trajectory tracking control [18].

The event-triggered control approach is further extended to the discrete-time case [24]-[26] where the sensor senses the system state periodically in a time triggered approach and the transmission of the feedback signals and controller execution are done at the event-triggering instants. A major advantage of the discrete time event-triggered control is that the minimum inter-event time is always guaranteed and is the periodic sampling interval of the discrete time system [26]. Similar to the event-triggered control in a discrete time domain, a periodic event-triggered control approach for continuous-time systems is presented in [7]. The triggering condition is evaluated periodically with a fixed sampling interval and the transmission decision is made at the violation of the condition. This design frame work enforces a positive lower bound on the minimum inter-event times. The stability analysis is carried out using three different modelling techniques used for hybrid systems such as impulsive system, piecewise linear systems, and perturbed linear systems. In all the above design approaches the system state or output and the control input are held between two consecutive events by a zero order hold (ZOH) for the implementation purpose.

In a second event-triggered approach [17], [24], [44], a model of the system is used to reconstruct the system state vector and, subsequently, used for designing the

control input. As the control input is based on the model states, no feedback transmission is required unless there is a significant change in the system performance due to external disturbance or internal parameter variation. In the area of model-based event-triggered control design, the authors in [17] used an input generator as a model to predict the system state and compute the control. Further, the authors in [44] consider the nominal dynamics of the system with uncertainty, usually of smaller magnitude and bounded, to form a model. The asymptotic stability was guaranteed by designing the event-triggering condition. A discrete-time model based approach is also presented in [24] for systems subjected to disturbance. Two modelling approaches (perturbed linear and piecewise linear system) are used to analyze the stability and global exponential stability with certain l_2 gain is guaranteed via linear matrix inequalities (LMI) based conditions. It is observed that the model-based approach reduces the number of events or transmissions more effectively when compared to the ZOH based approach, but, with a higher computational load due to induction of the model.

The ETC scheme is also extended to NCS with inherent network constraints [30]-[31], [34], [37], [40]-[41] such as constant or time varying delays, packet losses and quantization errors. In these design approaches, the event-triggering condition is tailored [30]-[31] to handle the maximum allowable delays, packet losses [30] and quantization error for both state and control input [37] so as to ensure stability. From optimal control point of view in the event-triggered context, a few results are available in the literature [21], [40]-[42]. Optimal event-triggered control for stochastic continuous time NCS is presented in [41]. The problem is formulated as an optimal stopping problem and an analytical solution is provided. The optimal control in an constrained networked

environment is studied in [42]. Further, the authors in [40] extended the work to event-triggered control frame work and characterized the certainty equivalence controller to be optimal in a linear quadratic Gaussian (LQG) frame work. The optimal control input and the optimal event-triggering instants are designed using the separation principle.

Despite these results from the literature on event-triggered control, all these schemes consider either the complete knowledge of the system dynamics [14]- [24], [26]-[39] or system with a smaller uncertainty [25], [44] with known nominal dynamics. In contrast, an L1 adaptive control scheme with known nominal dynamics is proposed in [43] where an adaptive law is used to estimate the uncertainty. Therefore, a comprehensive theory for the adaptive event-triggered control of a complete uncertain system is yet to be developed. Moreover, the optimal solution of the event-based control [21], [40] requires the system dynamics and backward in time solution of the Riccati equation making it difficult to implement. Thus, a forward in time and online solution to the optimal control problem in an event-triggered context is still an open problem.

In general, adaptive dynamic programming (ADP) [51], [59] and Q-learning [45], [50], [57] based schemes are used for a forward-in-time solution to the optimal control problems. The ADP was proposed by the authors in [51], [53], [59]-[60] and later became popular with various other names such as approximate dynamic programming (ADP) [60] and neuro-dynamic programming (NDP) [53]. These schemes in general use online approximator based parameterization and value and/or policy iterations [55], [60] to solve the Bellman or Hamilton-Jacobi-Bellman (HJB) equation to obtain the optimal control.

The policy iteration based techniques require a large number of iterations to maintain the stability [54]. Therefore, online implementation of these iterative schemes

are computational intensive and not practically viable. In contrast, [54], [61] proposed a time-based scheme to solve the ADP based optimal control in an on-line manner for discrete-time nonlinear systems. The time histories of the cost-to-go errors were used for the approximation of value function. A similar time-based technique is presented in [45] for linear networked control systems (NCS) in a stochastic framework by using Q-learning. In both the approaches, learning of the value function [54] or the Q-function [61] and controller executions were carried out periodically at every sampling instant. However, as mentioned earlier, the periodic sampled controller schemes will lead to higher cost for systems with limited computational and communication bandwidth resources. Therefore, an event sampled ADP and Q-learning scheme is needed for effective control of systems with sparse resources.

Motivated by the above facts, in this dissertation, a suite of novel event sampled adaptive control designs for uncertain linear and nonlinear systems is presented. The adaptive event sampled design is extended to event sampled optimal adaptive control schemes using ADP and Q-learning techniques with limited feedback information for systems with completely unknown dynamics. Adaptive and neural network based learning methods with intermittently available information are used to learn the unknown parameters/dynamics and a forward in time solution is presented. Lyapunov stability analysis is used to guarantee stability of the closed-loop event sampled systems. Next, the organization of the thesis is presented.

1.2 ORGANIZATION OF THE DISSERTATION

In this dissertation, event sampled adaptive regulation schemes of uncertain linear and nonlinear systems are developed. The proposed designs use event sampled

transmission of feedback data, parameter/NN weight update schemes, and controller execution to effectively utilize the available resources, such as communication network bandwidth and computational capability. The event sampling or transmission instants are determined using adaptive conditions designed by using Lyapunov stability theory. The dissertation consists of five papers and each paper portrays a sequential development of the research work as outlined in Figure 1.2. The first three papers present the event sampled stable and optimal adaptive regulator designs for discrete time systems both for linear and nonlinear systems with applications to NCS. The last two papers extended the event sampled paradigm to continuous time domain where event sampled stable and near optimal regulators for nonlinear continuous-time systems are designed.

The first paper presents a novel event sampled optimal adaptive state and output feedback control scheme for uncertain linear discrete-time systems. The infinite horizon optimal control for both the state and output feedback is solved by using the event sampled Q-learning and adaptive dynamic programming technique. The designs do not require the knowledge of system dynamics and compute the event sampled optimal control input in a forward in time and online manner without using any value/policy iterations. Further, the Q-function parameters are updated only at the event sampling instants with intermittently available state and control input vector. The asymptotic convergence of the system state vector and the parameter estimation error is proven by using Lyapunov analysis by designing novel adaptive event sampling condition for both the schemes. This adaptive event sampling condition guarantees the accuracy of the parameter convergence with reduced computation. The event sampled Q-learning scheme is applied to NCS represented as continuous time linear system with inherent time-

varying network induced delays and random packet losses. The randomness of the delays and packet losses leads to a stochastic time-varying discrete time system. Therefore, event driven Q-learning developed Paper I is analyzed in a stochastic frame work and asymptotic stability in the mean is guaranteed with reduced computation and communication. The results are placed in Appendix A.

On the other hand, in the second paper, a nonlinear discrete-time system in Brunovsky canonical form is considered where the dynamics are considered unknown. First, the universal approximation property of neural network (NN) is revisited in an event sampled context. The event sampled approximation is subsequently used to design an adaptive state estimator (SE) to approximate the system dynamics and estimate the state vector. The SE dynamics and state vector are utilized to obtain the control input, during any two event sampled instants. In this case the event sampling condition turns out to be a function of system state and NN weight estimates to facilitate approximation. Further, the event sampling condition uses a dead-zone operator to prevent the unnecessary triggering of events due to NN reconstruction error once the system state and the NN weights converge to the ultimate bound.

In the third paper, a more general class of nonlinear discrete-time affine system is considered and a novel technique to solve the finite horizon optimal control in an event sampled paradigm is proposed. This proposed approach uses event sampled NN-based identifier in conjunction with actor-critic NNs to solve the Hamilton-Jacobi-Bellman (HJB) equation online. Similar to other papers, the event sampling condition is made adaptive to ensure approximation accuracy and the ultimate boundedness (UB) of the closed-loop system. This event sampled ADP scheme in an infinite horizon frame work

is applied to nonlinear NCS with network induced time varying delays and packet losses. Since the NCS leads to stochastic time varying system, as discussed earlier, stochastic analysis is carried out for the actor critic frame work used in Paper III. Ultimate boundedness in the mean of the closed-loop event sampled NCS with potential saving in communication and computation is shown. The detailed stochastic design and results are included in Appendix B.

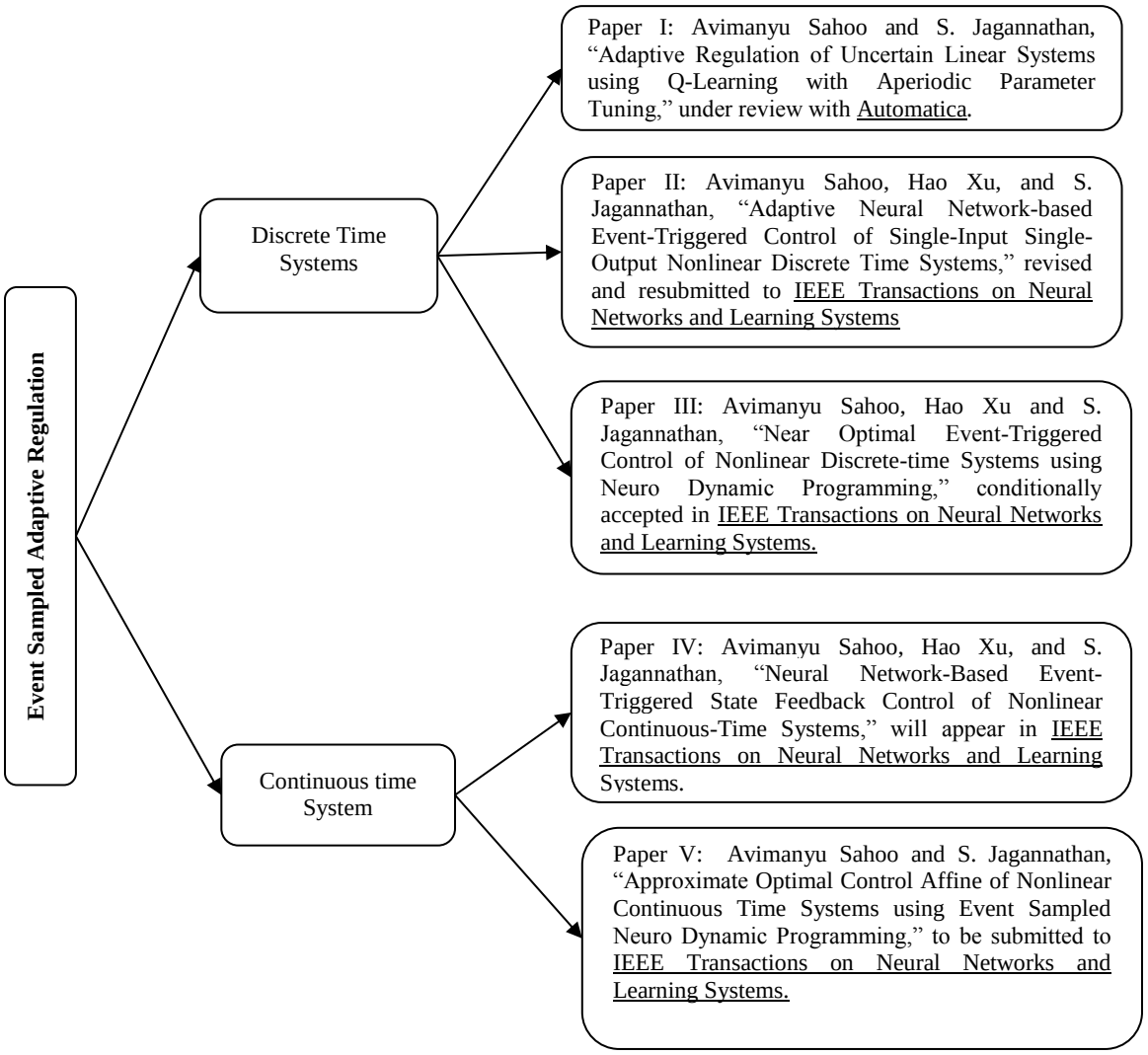

Figure 1.2. Organization of the dissertation.

In the last two papers, the event sampled designs of nonlinear continuous time dynamical systems are presented. A multi-input multi-output (MIMO) nonlinear affine system is considered in the fourth paper. An event sampled stabilizing control is developed using approximate feedback linearization. A linearly parameterized neural network is used to approximate the control input with event sampled feedback information. The event sampling condition is derived using the estimated neural network weights. The neural network weighs are updated as a jump at the event sampled instants and held during flow period. Therefore, the continuous time event sampled system is modeled as a nonlinear impulsive dynamical system to analyze the stability. The ultimate boundedness of the closed-loop system parameters are shown using extension of Lyapunov direct approach for impulsive dynamical systems. A positive lower bound on the inter-sample times is also guaranteed to avoid accumulation point.

In the final paper, the stabilizing controller design is extended to optimal control design by minimizing an infinite horizon cost function. Continuous time event sampled adaptive dynamic programming is developed to solve the optimal control problem with aperiodic sampled state and control input vectors. The value function, which is the solution of the Hamilton-Jacobi-Bellman equation, is approximated using neural networks and the weights are updated as a jump at the event sampled instants. A novel identifier is also designed with event sampled approximation of the neural network. The closed-loop stability is analyzed and ultimate boundedness is guaranteed by modelling the closed-loop system as in impulsive dynamical system. A formula for the minimum inter-sample time is derived to guarantee existence of the positive lower bound on the

inter-sample times to avoid Zeno behavior or accumulation point. The contributions of the dissertation are highlighted next.

1.3 CONTRIBUTIONS OF THE DISSERTATION

Traditional event-triggered control [13]-[44] as discussed in Section 1.1 are developed with the complete knowledge of the system dynamics [13]-[24], [26]-[42] or system with small bounded uncertainty [25], [44]. This made the problem rather simpler when compared to complete uncertain systems which are more pervasive in practical applications. Hence, event-based adaptive control schemes, where the controller adapts the changes in the system parameters in an online manner, will be of more practical importance. Therefore, the goal of this dissertation is to develop event-sampled regulation schemes for systems with completely uncertain system dynamics and at the same time retain the advantages of this approach in computation and communication saving. The contributions in this dissertation are summarized as follows.

The first contribution is the development of an event sampled optimal adaptive state feedback scheme for uncertain linear discrete time systems. However, the system state vector many not always be available for measurement and, hence, there is a need for the output feedback design. Therefore, an observer based optimal output feedback design is also provided. Traditional optimal control design needs a backward in time solution of Riccati equation [40], [48] or a forward in time iterative approach using policy and value iteration [51], [60] to solve the optimal control problem with periodic sampling scheme. In contrast, an event sampled Q-learning approach is developed with uncertain system dynamics for both state and output feedback. Parameter tuning is carried out at the event sampling instants only leading to an aperiodic update scheme to save computation when

compared to traditional adaptive control [56], [58]. The traditional event-triggering conditions developed for known system dynamics [13]-[40] are not suitable due to the adaptive nature of the proposed system. With this effect, novel adaptive event sampling conditions which not only guarantee the asymptotic convergence of system state but also the parameter estimation error to zero are designed. Asymptotic stability of the closed-loop system is demonstrated by using Lyapunov direct method. Further, application of this event sampled Q-learning scheme for NCS with network induced time varying delays and random packet losses is presented. A stochastic analysis of the event sampled Q-learning approach is presented to design the optimal control policy for NCS.

The main contribution of the second paper is the development of a neural network (NN) based event sampled adaptive regulation scheme for an uncertain nonlinear discrete-time system. The universal approximation property of the NN is revisited and an event sampled NN-based approximation is provided. In contrast with the model-based approach for system with known dynamics [24]-[25], an event sampled NN-based adaptive model design is presented. Further, aperiodic weight update scheme for the NN weights at the event sampling instants is proposed to save computation when compared to the traditional neural network based control. Moreover, the adaptive event sampling condition is designed to ensure the uniform ultimate boundedness of the closed-loop system.

The contributions of the third paper include the design of a finite horizon event sampled optimal control scheme for a more general class uncertain nonlinear discrete time system in affine form. Since the traditional time driven ADP schemes are computational intensive, an event sampled ADP design is provided. A novel neural

network based identifier is designed with event sampled availability of the feedback data to identify the system dynamics. The event sampled actor-critic frame work with time varying activation function to learn the finite horizon time varying value function is presented. Similar to the other papers, novel NN weight update schemes which tune the NN weights only at event sampling instants, to save computation, is proposed. The ultimate boundedness of the closed-loop system is demonstrated by using the Lyapunov technique and an adaptive event sampling condition. The application of this design approach for NCS is included in the appendix which leads to a stochastic event sampled ADP scheme. Finally, the ultimate boundedness for the event sampled system and ultimate boundedness in the mean of NCS is proved using Lyapunov stability analysis.

The contributions of the fourth paper include the design of an event sampled feedback linearized controller for continuous time nonlinear systems. The event sampled approximation in a continuous time domain is developed to approximate the control input using a linearly parameterized neural network. Non periodic update of the neural network weights as a jump at the event sampling instant is proposed. A nonlinear impulsive dynamical modelling of the event sampled system is presented. The adaptive event sampling condition is designed using the estimated neural network weights to decide the jump and flow periods. Ultimate boundedness of the system is also proved using extension of Lyapunov technique for impulsive dynamical systems.

The contributions of the last chapter include the design of an event sampled continuous-time adaptive dynamic programming based optimal controller. A novel event sampled identifier design is presented. The online approximation of the value function using event sampled HJB equation error is demonstrated. The impulsive modelling of the

event sampled near optimal system is included to analyze the system during the sampled instants and inter-sample times. The determination of the sampling condition in an optimal control frame work is proposed which guarantees accuracy of approximation. Finally, the ultimate boundedness of closed-loop parameters using the extension of Lyapunov technique, for impulsive dynamical system, is also included.

PAPER

I. ADAPTIVE REGULATION OF UNCERTAIN LINEAR SYSTEMS USING Q-LEARNING WITH APERIODIC PARAMETER TUNING

Avimanyu Sahoo and S. Jagannathan

***Abstract* —** *This paper presents a novel Q-learning based optimal adaptive state and output feedback control of uncertain linear discrete-time systems with aperiodic event-based feedback information. Both dynamic programming (DP) and Q-learning techniques with event sampled system and observer state vectors are used to design and learn the optimal control input sequence. The event-based time history of the temporal difference error in Bellman equation is utilized to find the solution to the Bellman equation in DP without using traditional policy and/or value iterations. The event sampled instants are determined via a trigger condition which is analytically derived by using Lyapunov stability criterion. The Q-function parameters are tuned only at the event sampled instants thereby leading to non-periodic parameter tuning. It is further shown that the closed-loop parameters converge asymptotically provided persistency of excitation condition on the control input is ensured. Simulation results are included to validate both the analytical designs. The net result is the development of event-driven dynamic programming via Q-learning for linear systems.*

***Keywords* —** Q-learning, event sampled adaptive dynamic programming, adaptive observer, aperiodic sampling, and optimal control.

1. INTRODUCTION

Optimal control (Lewis & Syrmos, 1995) is a key area of research among the control researchers in the past several decades. Adaptive dynamic programming (ADP) (Barto, Sutton, & Anderson, 1983, Watkins, 1989) drew more attention because of the forward-in-time solution to the optimal control problems. The ADP was proposed by Werbos (1992), Barto, Sutton, and Anderson (1983), Watkins (1989), Bertsekas and Tsitsiklis (1996), and later became popular with various other names such as approximate dynamic programming (ADP) (Werbos, 1992) and neuro-dynamic programming (NDP) (Bertsekas & Tsitsiklis, 1996). These schemes in general use online approximator based parameterization and value and/or policy iterations to solve the Bellman or Hamilton-Jacobi-Bellman (HJB) equation to obtain the optimal control in a forward-in-time manner (Wang, Jin, Liu, & Wei, 2011).

Among the ADP based Q-learning schemes, Bradtke, Ydestie, and Barto (1994) proposed policy iteration based adaptive Q-learning approach by using the system dynamics. In Hagen and Krose (1998), two approximation techniques were proposed to compute the optimal policy. The first used a model to identify the system dynamics and compute the Riccati solution, whereas, the second scheme used Q-learning with least square update. Later, the Q-learning scheme is extended by Tamimi, Lewis, and Murad (2007) to the zero-sum-game formulation by using model-free policy iteration.

The policy iteration based techniques require a large number of iterations to maintain the stability (Dierks & Jagannathan, 2012). Therefore, online implementation of these iterative schemes are computational intensive and not practically viable. In contrast, Dierks and Jagannathan (2012) proposed a time-based scheme to solve the ADP based

optimal control in an on-line manner. The time histories of the cost-to-go errors were used for the approximation of value function. A similar time-based technique has been used by Xu, Jagannathan, and Lewis (2012) for networked control systems (NCS) in a stochastic framework by using Q-learning. In both the approaches, learning of the value function (Dierks & Jagannathan, 2012) or the Q-function (Xu, Jagannathan, & Lewis, 2012) and controller executions were carried out periodically at every sampling instant. However, the periodic sampled controller schemes will lead to higher cost for systems with limited computational and communication bandwidth resources.

Recently, it was demonstrated that the state or event-based sampling and controller execution are advantageous over periodic time-driven sampling counterpart in terms of computational cost (Tabuada, 2007; Wang & Lemmon, 2011). Control design by using event-based sampling is referred to as event-triggered control (ETC) (Donkers & Heemels, 2012; Tabuada, 2007; Wang & Lemmon, 2011). The aperiodic event-based sampling instants are determined by using a triggering condition while maintaining stability and performance. This event-triggering condition uses the state or output information (Donkers & Heemels, 2012; Tabuada, 2007; Wang & Lemmon, 2011) given the system dynamics.

The traditional optimal control (Lewis & Syrmos, 1995) in the context of limited communication and event-based sampling is studied by Imer and Basar (2006), Cogill (2009), Molin and Hirche (2013) and others. A backward-in-time solution of the Riccati equation with separation principle is being used. To the best knowledge of authors, this is the first time a forward-in-time and online optimal control scheme using ADP and Q-learning technique with event sampled state information for uncertain linear systems is

presented. In a traditional discrete time system framework, the term sampling represents the periodic time instants with a fixed sampling interval. By contrast, in this paper, the term event sampled instant refers to the aperiodic time instants when the feedback signals are made available to the controller and the parameters are tuned by using a triggering condition.

Event-based sampling requires the development of event-driven dynamic programming (DP) and a redesign of the controller with temporal difference (TD) or Bellman error. Therefore, in the first part of this paper, a novel Q-learning based optimal state feedback control scheme, for uncertain linear discrete-time systems, with event-based sampling, is introduced. However, since the state vector is unavailable for measurement in many applications, an output feedback optimal design using an adaptive observer is also presented next. A Q-function estimator (QFE) is used to learn the optimal action dependent value or Q-function on-line for both state and output feedback cases at event sampled instants.

The time-histories of the Bellman errors from Bellman equation are used to tune the QFE parameters at the event sampled instants and, hence, the parameters are tuned in an aperiodic manner. The control input is, subsequently, updated from the QFE parameters at event sampled instants. This aperiodic tuning saves the computations when compared to traditional adaptive Q-learning. Above all, the adaptive triggering conditions to determine the event sampled instants are derived analytically. These conditions ensure the convergence of parameters by creating a sufficient number of event sampled instants during the initial adaptation while keeping the computation small. Finally, the stability of the event sampled closed-loop system was demonstrated by using the Lyapunov method

(Wang & Lemmon, 2011). A preliminary version of the work in a finite horizon optimal control frame work is published as Sahoo and Jagannathan (2014). Here optimization of the event sampled instants is not considered.

Thus, the primary **contributions** of this paper include: 1) the development of optimal adaptive state and output regulation schemes for uncertain linear systems with event-based sampling, 2) the design of tuning scheme for online estimation of QFE parameters, for both state and output feedback, 3) the development of adaptive triggering conditions, and 4) the demonstration of closed-loop stability in the presence of uncertain dynamics and aperiodic sampling. The next section presents a brief background on the traditional Q-learning for both state and output feedback.

2. BACKGROUND

After a brief Q-learning background, the optimal control problem statement with event based sampling is introduced.

2.1 STATEFEEDBACK DESIGN

Consider the linear time-invariant (LTI) discrete-time system given as

$$x_{k+1}^p = Ax_k^p + Bu_k,\ y_k^p = Cx_k^p, \tag{1}$$

where $x_k^p \in \Omega_x \subset \Re^n$, $y_k^p \in \Omega_y \subset \Re^{p_o}$, and $u_k \in \Omega_u \subset \Re^m$ represent the system state, the output and the control input vectors, respectively. The system matrices, $A \in \Re^{n\times n}$ and $B \in \Re^{n\times m}$, are considered unknown. The output matrix $C \in \Re^{p_o \times n}$ is known. The system (1) satisfies the following assumption.

Assumption 2.1. The system (1) is considered controllable and observable with the control coefficient matrix satisfying $\|B\| \le B_{\max}$, where $B_{\max} > 0$ is a known constant. Further, the order of the system is known.

Consider the value function for (1) given by

$$V_k = \sum_{j=k}^{\infty} r(x_j^p, u_j), \tag{2}$$

where $r(x_k^p, u_k) = x_k^{p^T} P x_k^p + u_k^T R u_k$ is a positive definite cost-to-go function at the time instant k. The matrices $P \in \Re^{n\times n}$ and $R \in \Re^{m\times m}$ are, respectively, positive semi-definite and definite matrices to penalize the system state x_k^p and the control input u_k. The initial control input u_0 is assumed to be admissible to keep the value function (2) finite. Traditionally the sequence of control inputs, u_k, which minimizes the value function (2)

can be obtained by solving the algebraic Riccati equation (ARE) (Lewis & Syrmos, 1995).

The solution to the ARE, for computing the optimal control input, is not feasible when the system dynamics A and B are not known. Adaptive Q-learning based techniques (Tamimi, Lewis, & Murad, 2007; Xu, Jagannathan, & Lewis, 2012) on the other hand are employed to generate optimal control input sequence without using system dynamics.

Define the action dependent value or the Q-function (Bradtke, Ydestie, & Barto, 1994; Tamimi, Lewis, & Murad, 2007; Xu, Jagannathan, & Lewis, 2012) as

$$Q^*(x_k^p, u_k) = r(x_k^p, u_k) + V_{k+1}^* = [x_k^{p^T} \quad u_k^T] G [x_k^{p^T} \quad u_k^T]^T, \tag{3}$$

where $G = \begin{bmatrix} P + A^T SA & A^T SB \\ B^T SA & R + B^T SB \end{bmatrix} \equiv \begin{bmatrix} G^{xx} & G^{xu} \\ G^{ux} & G^{uu} \end{bmatrix}$ with V_{k+1}^* being the optimal value function from time instant $k+1$ onwards. The optimal control input using the Q-function (3) is written as $u_k^* = -K^* x_k^p$ where $K^* = (R + B^T SB)^{-1} B^T SA = (G^{uu})^{-1} G^{ux}$. Therefore, the optimal control input sequence can be computed online in a forward-in-time manner by estimating the Q-function (3). The *Q*-function (3) in parametric form is given by

$$Q^*(x_k^p, u_k) = z_k^{p^T} G z_k^p = \Theta^T \xi_k^p, \tag{4}$$

where $z_k^p = [x_k^{p^T} \quad u_k^T]^T \in \Re^l$ with $l = m + n$, $\xi_k^p = z_k^p \otimes z_k^p$ is a quadratic polynomial or regression vector, $\otimes$ denotes the Kronecker product, and $\Theta \in \Omega_\Theta \subset \Re^{l_g}$ is the Q-function parameter vector formed by vectorization of the parameter matrix G with $l_g = l(l+1)/2$, as given in Xu, Jagannathan, and Lewis (2012).

The estimate of the Q-function (4), $\hat{Q}(x_k^p, u_k)$, by the QFE with periodically sampled state vector is expressed as

$$\hat{Q}(x_k^p, u_k) = z_k^{p^T} \hat{G}_k z_k^p = \hat{\Theta}_k^T \xi_k^p, \tag{5}$$

where $\hat{\Theta}_k \in \Re^{l_g}$ is the estimate of Q-function parameter vector Θ referred to as QFE parameters and $\hat{G}_k = \begin{bmatrix} \hat{G}_k^{xx} & \hat{G}_k^{xu} \\ \hat{G}_k^{ux} & \hat{G}_k^{uu} \end{bmatrix}$ represents the estimation of G. The Q-function is equal to the optimal value function V_k^* when the control input is optimal. Thus, we have $V_k^* = \min_{u_k} Q^*(x_k^p, u_k)$. By Bellman's principle of optimality, the optimal value function satisfies

$$0 = V_{k+1}^* - V_k^* + r(x_k^p, u_k) = r(x_k^p, u_k) + \Theta^T \Delta \xi_k^p, \tag{6}$$

where $\Delta \xi_k^p = \xi_{k+1}^p - \xi_k^p$. Since the estimated Q-function (5) does not satisfy (6), the temporal difference (TD) error or the Bellman error is given by

$$e_k^V = r(x_k^p, u_k) + \hat{\Theta}_k^T \Delta \xi_k^p. \tag{7}$$

Instead of using iteration based techniques (Tamimi, Lewis, & Murad, 2007), it has been shown by Xu, Jagannathan and Lewis, (2012) that an optimal control input can be obtained by tuning QFE parameter $\hat{\Theta}_k$ with the time history of the Bellman error in an forward-in-time and online manner. Next, the optimal control using output feedback is introduced.

2.2 OUTPUT FEEDBACK DESIGN

The value function for (1) using the output can be redefined as $V_k = \sum_{j=k}^{\infty} r^y(y_j^p, u_j)$ where $r^y(y_j^p, u_j) = y_j^{p^T} P^y y_j^p + u_j^T R u_j$ is the cost-to-go and

$P^y \in \Re^{p_o \times p_o}$ is a positive definite matrix to penalize system output. The value function can be rewritten by using the output equation in (1) as

$$V_k = \sum_{j=k}^{\infty} x_j^{p^T} C^T P^y C x_j^P + u_j^T R u_j = \sum_{j=k}^{\infty} x_j^{p^T} P x_j^P + u_j^T R u_j \,, \tag{8}$$

where $P = C^T P^y C$. From (8), the state feedback Q-learning based design can be utilized to the output feedback case by reconstructing the system state vector. An adaptive observer similar to the one from Zhao, Xu, and Jagannathan (2014) can be used to reconstruct the system state and is given next.

Consider the following adaptive observer dynamics

$$x_{k+1}^o = \hat{A}_k x_k^o + \hat{B}_k u_k + \hat{L}_k (y_k^p - y_k^o), \;\; y_k^o = C x_k^o, \tag{9}$$

where $x_k^o \in \Re^n$ and $y_k^o \in \Re^{p_o}$ represent the observer state and the output vectors, $\hat{A}_k \in \Re^{n \times n}$ and $\hat{B}_k \in \Re^{n \times m}$ denote the estimated observer system matrices and $\hat{L}_k \in \Re^{n \times p_o}$ is the estimated observer gain matrix. The observer matrices are estimated using a parametric form given by

$$x_{k+1}^o = \hat{\psi}_k^T \phi_k \,, \tag{10}$$

where $\hat{\psi}_k = [\hat{A}_k \quad \hat{B}_k \quad \hat{L}_k]^T \in \Re^{q \times n}$ is the estimated observer parameter matrix, $\phi_k = [x_k^{o^T} \quad u_k^T \quad e_k^{y^T}]^T \in \Re^q$ is the regression vector, $e_k^y = y_k^p - y_k^o$ is the observer output error, and $q = n + m + p_o$. The estimated observer parameter matrix $\hat{\psi}_k$ is tuned at every sampling instant $k \in \mathbb{N}$ so that the state estimation error given by $e_k^x = x_k^p - x_k^o$ converges to zero.

The event sampled system and observer state vector require a redesign of the optimal controller as discussed next.

3. PROBLEM STATEMENT

In this section, the event sampled optimal control problem is formulated. The estimation and stability issues with event sampled state vector and parameter tuning are addressed for both state and output feedback design.

3.1 STATE FEEDBACK DESIGN

The structure of an event-based Q-learning scheme is illustrated in Figure 1. Here, the system state vector x_k^p and the control input u_k are sent to the controller and the plant, respectively, only at the event sampling instants (when the switches are in closed position).

Let the subsequence k_i, $i=1,2,\cdots$, of $k \in \mathbb{N}$ represent the event sampling instants with $k_0 = 0$ being the initial sampling instant. The system state vector $x_{k_i}^p$, $i=1,2,\cdots$ sent to the controller, is held by a zero order hold (ZOH) until the next sampling instant and it is expressed as

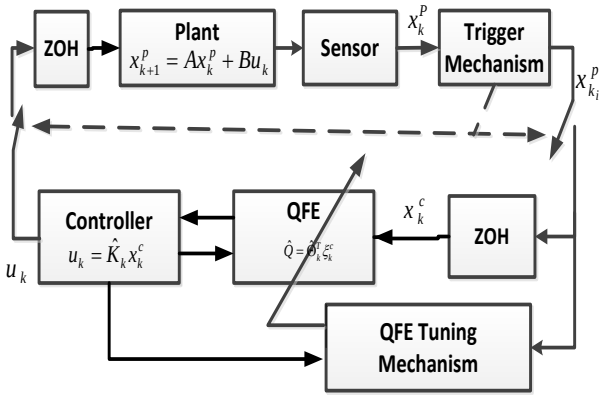

Figure 1. Q-learning based optimal feedback regulator with aperiodic update.

$$x_k^c = \begin{cases} x_k^p, & k = k_i, \\ x_{k_i}^p, & k_i < k < k_{i+1}, \end{cases} \tag{11}$$

where x_k^c is the last held state at the ZOH. The error between the current system state vector, x_k^p, and the last held system state vector, x_k^c, at the ZOH is referred to as event sampling error and it is given by

$$e_k^{ET} = x_k^p - x_k^c,\ k_i \le k < k_{i+1},\ i = 1, 2, \cdots. \tag{12}$$

The event sampled instants are decided by comparing the event sampling error with a state dependent threshold (to be computed later) (Donkers & Heemels, 2012; Tabuada, 2007) referred to as triggering condition. The triggering condition is evaluated at every periodic time instant k at the trigger mechanism and a decision is made whether or not to release the system state vector. The system state vector is released at the violation of the condition. Upon receiving the system states at the ZOH, the last held state is reset to the current measured value as in (11) and the event sampling error (12) is reset to zero for the next event.

Our objective is to design an optimal controller by minimizing (2) with event sampled state vector. The event sampled optimal control input sequence for the cost function (2) when used with a Q-function can be written as

$$u_k^* = -K_k^* x_k^c = -(G^{uu})^{-1} G^{ux} (x_k^p + e_k^{ET}),\ k_i \le k < k_{i+1}. \tag{13}$$

This optimal control input (13) is governed by e_k^{ET} and the estimation of G using event sampled state vector. The Q-learning approach in Section 2 cannot be utilized directly to estimate the QFE parameter $\hat{\Theta}_k$ or $\hat{G}_k$ with event sampled state vector due the following reasons.

For the Q-function parametric form given in (4), the Q-function estimate by using the event sampled state vector (11) , in contrast with (5), is given by

$$\hat{Q}(x_k^c, u_k) = z_k^{c^T} \hat{G}_k z_k^c = \hat{\Theta}_k^T \xi_k^c, \quad k_i \le k < k_{i+1} , \tag{14}$$

where $z_k^c = [x_k^{c^T} \quad u_k^T]^T$ and $\xi_k^c = z_k^c \otimes z_k^c$ being the event sampled regression vector. The Bellman error with event sampled state can be represented as

$$e_k^V = r(x_k^c, u_k) + \hat{\Theta}_k^T \Delta\xi_k^c, \; k_i \le k < k_{i+1}, \tag{15}$$

where $r(x_k^c, u_k) = x_k^{c^T} P x_k^c + u_k^T R u_k$ and $\Delta\xi_k^c = \xi_{k+1}^c - \xi_k^c$. The Bellman error (15) in terms of the periodic system state is rewritten as

$$e_k^V = r(x_k^p, u_k) + \hat{\Theta}_k^T \Delta\xi_k^p + \Xi_s\left(x_k^p, e_k^{ET}, \hat{\Theta}_k\right), \tag{16}$$

where $\Xi(x_k^p, e_k^{ET}, \hat{\Theta}_k) = r((x_k^p - e_k^{ET}), u_k) - r(x_k^p, u_k) + \hat{\Theta}_k^T(\Delta\xi_k^c - \Delta\xi_k^p)$. By comparing the event sampled Bellman error (16) with the periodic sampled one from (7), the error (16) includes an additional error $\Xi_s(x_k^p, e_k^{ET}, \hat{\Theta}_k)$. This additional error consists of errors in cost-to-go, $(r((x_k^p - e_k^{ET}), u_k) - r(x_k^p, u_k))$ and the regression vector, $\hat{\Theta}_k^T(\Delta\xi_k^c - \Delta\xi_k^p)$ which are driven by e_k^{ET}. Hence, the accuracy of the estimation of QFE parameters depends upon the threshold for the event sampling error in the triggering condition. A smaller threshold value will limit the event sampling error and will ensure a better accuracy. On the other hand, this will lead to more events and in turn higher computation. Therefore, a trade-off has to be reached via a suitable triggering condition design.

3.2 OUTPUT FEEDBACK DESIGN

In this case the observer states are sent to the controller at the event sampled instants. The event sampled observer state can be defined as

$$x_k^{o,c} = \begin{cases} x_k^o, & k=k_i, \\ x_{k_i}^o, & k_i < k < k_{i+1}, \end{cases} \tag{17}$$

where $x_k^{o,c}$ is last held observer state at the ZOH. The observer based event sampling error can be redefined as

$$e_k^{o,ET} = x_k^o - x_k^{o,c},\ k_i \le k < k_{i+1}. \tag{18}$$

Now to estimate the Q-function the event sampled observer state vector is used. The QFE by using (17) can be written as

$$\hat{Q}(x_k^{o,c}, u_k) = z_k^{o,c^T} \hat{G}_k z_k^{o,c} = \hat{\Theta}_k^T \xi_k^{o,c},\ \ k_i \le k < k_{i+1}, \tag{19}$$

where $z_k^{o,c} = [x_k^{o,c^T} \quad u_k^T]^T$ is the event sampled observer based regression vector.

Similar to the state feedback case, the Bellman error using event sampled observer state vector (17) is given by

$$e_k^{o,V} = r(x_k^{o,c}, u_k) + \hat{\Theta}_k^T \Delta\xi_k^{o,c},\ \ k_i \le k < k_{i+1}, \tag{20}$$

where $r(x_k^{o,c}, u_k) = x_k^{o,c^T} P x_k^{o,c} + u_k^T R u_k$ and $\Delta\xi_k^{o,c} = \xi_{k+1}^{o,c} - \xi_k^{o,c}$.

This event sampled Bellman error can be expressed in terms of the periodic system state as

$$e_k^{o,V} = r(x_k^p, u_k) + \hat{\Theta}_k^T \Delta\xi_k^p + \Xi_o(x_k^p, e_k^x, e_k^{ET}, \hat{\Theta}_k), \tag{21}$$

where $\Xi_o(x_k^p, e_k^x, e_k^{ET}, \hat{\Theta}_k) = r((x_k^p - e_k^x - e_k^{ET}), u_k) - r(x_k^p, u_k) + \hat{\Theta}_k^T(\Delta\xi_k^c - \Delta\xi_k^p)$. Similar to the state feedback case, the event sampled Bellman error is a function of observer event

sampling error $e_k^{o,ET}$ and state estimation e_k^x. Therefore, for the observer based design both a suitable triggering condition and update law to tune the observer parameters are needed.

Further, due to the availability of the system and observer state vectors at the event sampling instants alone, the QFE parameters must be tuned at these aperiodic instants. The frequency of the parameter tuning is a function of the event sampling error and, hence, the triggering condition. From the above discussion, unlike the traditional event-triggered control (Tabuada, 2007), the design of the triggering condition should not only ensure the stability with a reduction in computation but also facilitate the estimation of QFE parameters with event sampled system and observer states. This makes the design involved especially the triggering condition and proof of stability. Next, the above mentioned issues are mitigated by using a novel design framework.

4. EVENT SAMPLED STATE FEEDBACK DESIGN

4.1 PROPOSED SOLUTION

In this paper, we propose an adaptive threshold for the triggering condition by using the estimated QFE parameters and system state vector given in (29). This orchestrates the estimation process along with reduction in computation and discussed in details in Remark 4.5. The Q-function parameters are estimated locally at the trigger mechanism by a mirror QFE to evaluate the adaptive triggering condition. This saves transmissions of Q-function estimated parameters from the QFE at the controller to the trigger mechanism in case of an NCS. The mirror and the actual QFE operate in synchronism and initialized with same initial conditions. Note that, the addition of a mirror QFE increases the computational load. But, the overall computation is reduced due to aperiodic execution of the control input and QFE parameter tuning law both at mirror and controller. The time histories of the event sampled Bellman error is used in the tuning law. Further, as we wil see the previous state $x^p_{k_i-1}$ in addition to the current state $x^p_{k_i}$ is required for QFE parameter tuning, we propose to send both the states together to the controller at the event sampled instants. More discussion is given in Remark 4.1.

4.2 CONTROLLER DESIGN AND APERIODIC LEARNING OF Q-FUNCTION

Recall the event sampled QFE given in (14). The estimated control gain matrix now can be obtained from the QFE estimated parameter vector $\hat{\Theta}_k$ or $\hat{G}_k$ in (14). In terms of the estimated parameters, the control input is given by

$$u_k = -\hat{K}_k x_k^c, \;\; k_i \le k < k_{i+1}, \tag{22}$$

where $\hat{K}_k = (\hat{G}_k^{uu})^{-1}\hat{G}_k^{ux}$ is the estimated control gain.

The QFE parameter vector $\hat{\Theta}_k$ is tuned by using the history of the Bellman error (15) that is available at the event sampling instants. By using this, the auxiliary Bellman error at the event sampling instants $k = k_i$ is expressed as

$$\Xi_k^V = \Pi_k^p + \hat{\Theta}_k^T Z_k^p, \;\; k = k_i, \tag{23}$$

where $\Pi_k^p = [r(x_{k_i}^p, u_{k_i}) \;\; r(x_{k_{i-1}}^p, u_{k_{i-1}}) \;\; \cdots \;\; r(x_{k_{i-1-j}}^p, u_{k_{i-1-j}}] \in \Re^{1\times j}$ and $Z_k^p = [\Delta\xi_{k_i}^p \;\; \Delta\xi_{k_{i-1}}^p \;\; \cdots \;\; \Delta\xi_{k_{i-1-j}}^p] \in \Re^{l_g \times j}$ for $k = k_i$ with $0 < j < i$. The auxiliary Bellman error (23) uses the current estimated QFE parameter vector $\hat{\Theta}_k$ to evaluate the error. This makes the learning faster. The number of previous value j depends upon past experience and a value $j \le l_g$ is found suitable during simulation studies. A larger time history may lead to a faster convergence whereas it leads to higher computation.

Next update law for the QFE estimated parameter vector $\hat{\Theta}_k$, tuned only at the event sampling instants, is given by

$$\hat{\Theta}_k = \begin{cases} \hat{\Theta}_{k-1} - \dfrac{W_{k-1} Z_{k-1}^p \Xi_{k-1}^{V^T}}{\left\| I + Z_{k-1}^{p^T} W_{k-1} Z_{k-1}^p \right\|}, & k = k_i, \\ \hat{\Theta}_{k-1}, & k_{i-1} < k < k_i, \end{cases} \tag{24}$$

where

$$W_k = \begin{cases} W_{k-1} - \dfrac{W_{k-1} Z_{k-1}^p Z_{k-1}^{p^T} W_{k-1}}{\left\| I + Z_{k-1}^{p^T} W_{k-1} Z_{k-1}^p \right\|}, & k = k_i \\ W_{k-1}, & k_{i-1} < k < k_i \end{cases}, \tag{25}$$

with $W_0 = \beta I$, $\beta > 0$ a large positive value and I is the identity matrix of appropriate dimension. Because of the aperiodic execution of (24), it saves computation when compared to the traditional adaptive Q-learning techniques (Xu, Jagannathan, & Lewis, 2012).

Remark 4.1. The QFE parameter tuning law (24) requires the state vectors $x_{k_i}^p$ and $x_{k_i-1}^p$ for the computation of Z_{k-1}^p at $k = k_i$. Thus, both the state vectors are sent to the controller together at the event sampling instants as proposed.

Defining the QFE parameter estimation error $\tilde{\Theta}_k = \Theta - \hat{\Theta}_k$, the error dynamics using (24) , can be represented as

$$\tilde{\Theta}_{k+1} = \begin{cases} \tilde{\Theta}_k + (W_k Z_k^p \Xi_k^{V^T} \Big/ \left\| I + Z_k^{p^T} W_k Z_k^p \right\|), \ k = k_i, \\ \tilde{\Theta}_k, \ \ k_i < k < k_{i+1}. \end{cases} \tag{26}$$

Remark 4.2. The QFE parameter estimation error $\tilde{\Theta}_k$ will converge to zero if the augmented matrix Z_k^p satisfies the persistency of excitation (PE) condition (Green, & Moore, 1986). This can be achieved by ensuring the the regression vector ξ_k^p satisfies PE condition. The definition of the PE condition is presented next for completeness.

Definition 4.3. (Goodwin & Sin, 1984) A vector $\varphi(x_k)$ is said to be persistently exciting over an interval if there exist positive constants δ , $\underline{\alpha}$, $\bar{\alpha}$, and $k_d \geq 1$ such that $\underline{\alpha} I \leq \sum_{k=k_d}^{k+\delta} \varphi(x_k)\varphi^T(x_k) \leq \bar{\alpha} I$, where I is the identity matrix of appropriate dimension.

A PE like condition for the regression vector ξ_k^p can be achieved by adding an exploration noise to the control input u_k during the estimation process (Xu, Jagannathan, & Lewis, 2012).

The Bellman error e_k^V at the sampling instants in terms of $\tilde{\Theta}_k$ can be computed by subtracting (6) from (15) with $x_k^c = x_k^p$ for $k = k_i$. It is given by $e_k^V = -\tilde{\Theta}_k^T \Delta\xi_k^p,\ k = k_i$. Thus, the auxiliary Bellman error (23) at the event sampling instant becomes

$$\Xi_k^V = -\tilde{\Theta}_k^T Z_k^p,\ k = k_i. \tag{27}$$

This expression will be used for proving the asymptotic convergence of the QFE parameter estimation error presented next.

Lemma 4.4. Consider both the QFE (14) and the tuning law (24) and (25) with an initial admissible control policy $u_0 \in \Re^m$. Let the Assumption 2.1 holds and the QFE parameter vector $\hat{\Theta}_0$ be initialized to be non-zero in a compact set Ω_Θ. Under the assumption that the regression vector ξ_k^p satisfies PE condition, there exists a constant $\beta > 0$ such that the QFE parameter estimation error $\tilde{\Theta}_k$ converges to zero asymptotically when the event sampling instants $k_i \to \infty$ or, alternatively, $k \to \infty$.

Proof. Refer to Appendix.

4.3 TRIGGER CONDITION AND CONVERGENCE

The system dynamics (1) and the estimated control input (22) can be used to represent the closed-loop system dynamics as

$$x_{k+1}^p = Ax_k^p - B\hat{K}_k x_k^p + B\hat{K}_k e_k^{ET},\ \ k \le k < k_{i+1}. \tag{28}$$

To ensure the closed-loop stability of the system and reduction in computation along with estimation of the QFE parameters, the following criterion given by

$$\left\| e_k^{ET} \right\| \leq \sigma_k^{ETC} \left\| x_k^p \right\|, \tag{29}$$

is selected as the trigger condition where $\sigma_k^{ETC} \left\| x_k^p \right\|$ is the adaptive threshold, $\sigma_k^{ETC} = \sqrt{\Gamma(1-3\mu) \Big/ 3B_{\max}^2 \left\| \hat{K}_k \right\|^2}$ being the threshold coefficient, with $0 < \Gamma < 1$, and $0 < \mu < 1/3$. To ensure $\left\| \hat{K}_k \right\|$ in the threshold coefficient non zero while evaluating the triggering condition, the previous non-zero value for $\left\| \hat{K}_k \right\|$ is used when the estimated control gain becomes zero. The event sampled instants are decided at the violation of the condition (29).

Remark 4.5. The threshold coefficient σ_k^{ETC} uses the estimated control gain matrix $\hat{K}_k$ computed from the QFE parameter estimate vector $\hat{\Theta}_k$. Thus, the trigger condition (29) is adaptive in nature and is implicitly driven by the QFE parameter error $\tilde{\Theta}_k$. This facilitates the learning of the Q-function parameters by generating required event sampled instants. Once the QFE parameters converge to their target values the threshold coefficient becomes constant which is same as the traditional event-trigger condition (Tabuada, 2007). This further implies that for different initial values of $\hat{\Theta}_0$ in (24) and W_0 in (25), the threshold will be adjusted accordingly to generate required number of events during the initial adaption phase.

The following lemma is necessary before the main results can be claimed.

Lemma 4.6. Consider the controllable linear discrete-time system given by (1). Then there exists an optimal control input sequence, u_k^*, such that the closed-loop dynamics are expressed as

$$\left\| Ax_k^p + Bu_k^* \right\|^2 \le \mu \left\| x_k^p \right\|^2, \tag{30}$$

where $0 < \mu \le 1$ is a constant.

Proof. Consider the Lyapunov function $L(x_k^p) = x_k^{p^T} x_k^p$. The first difference along the system dynamics (1) is expressed as $\Delta L(x_k^p) \le \left\| Ax_k^p + Bu_k^* \right\|^2 - \left\| x_k^p \right\|^2 \le -(1-\mu)\left\| x_k^p \right\|^2$.

Since, the system is controllable and the optimal control input u_k^* is stabilizing (Lewis & Syrmos, 1995), the first difference $\Delta L\left(x_k\right) \le 0$. This implies the parameter μ should satisfy $0 < \mu \le 1$. ■

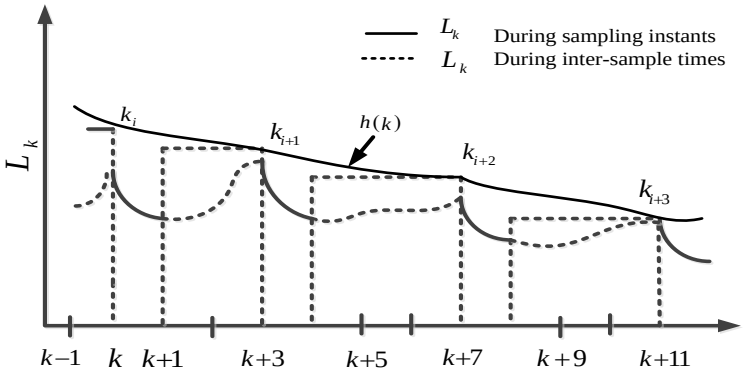

Figure 2. Evolution of the Lyapunov function at event sampling instants and inter-event times.

Next, the asymptotic stability of the closed-loop system is shown by evaluating a single Lyapunov function L_k^s both during the event sampling instants and the inter-event times as shown in Figure 2. It is shown that, the Lyapunov function is not decreasing monotonically during both the cases. This is also not necessary to prove stability of event sampled systems as discussed by Wang and Lemmon (2011). The Lyapunov function

may increase during the inter-event times. We only need to show the existence of a piecewise continuous function $h(k) \in \Re^+$, such that

$$h(k) \geq L_k^s \text{. for all } k \in \mathbb{N} \text{ and} \lim_{k \to \infty} h(k) = 0 \text{.} \tag{31}$$

Theorem 4.7. Consider the closed-loop system (28), QFE (14), QFE parameter estimation error dynamics (26) along with the control input (22). Let the Assumption 2.1 holds, $u_0 \in \Omega_u$ be an initial admissible control policy and the regression vector ξ_k^p satisfies PE condition. Suppose the last held state vector, x_k^c, and the QFE parameter vector, $\hat{\Theta}_k$ are updated by using (11), (24) and (25), respectively, at the violation of the triggering condition (29). Then, there exists a constant $\beta > 0$ such that the closed-loop system state vector x_k^p for all $x_0^p \in \Omega_x$ and the QFE parameter estimation error $\tilde{\Theta}_k$ for all non-zero $\hat{\Theta}_0 \in \Omega_\Theta$ converge to zero asymptotically with event sampling instants $k_i \to \infty$ or, alternatively, time instants $k \to \infty$. Further, the estimated Q-function $\hat{Q}(x_k^p, u_k) \to Q^*(x_k^p, u_k^*)$ and estimated control input $u_k \to u_k^*$ as $k \to \infty$.

Proof. Refer to the Appendix.

The flowchart shown in Figure 3 illustrates the implementation of the scheme. Since, the initial event sampling instant is considered at k_0, the initial system state and the state held by the ZOH are initialized with same value. The Q-function parameters both at the trigger mechanism and controller are also initialized with the same value. The system is operated with the initial control input. Then, the triggering condition is evaluated and the decision for releasing the system state is made if the event sampling error is greater than or equal to the threshold. The QFE both at the trigger mechanism and the controller

gets updated. Next, the control input is updated and sent to the plant and the time is incremented. If the event sampling error is below threshold the QFE and control input are not updated and time is incremented for next iteration. In the next section, the state feedback design is extended to an output feedback case using an adaptive observer.

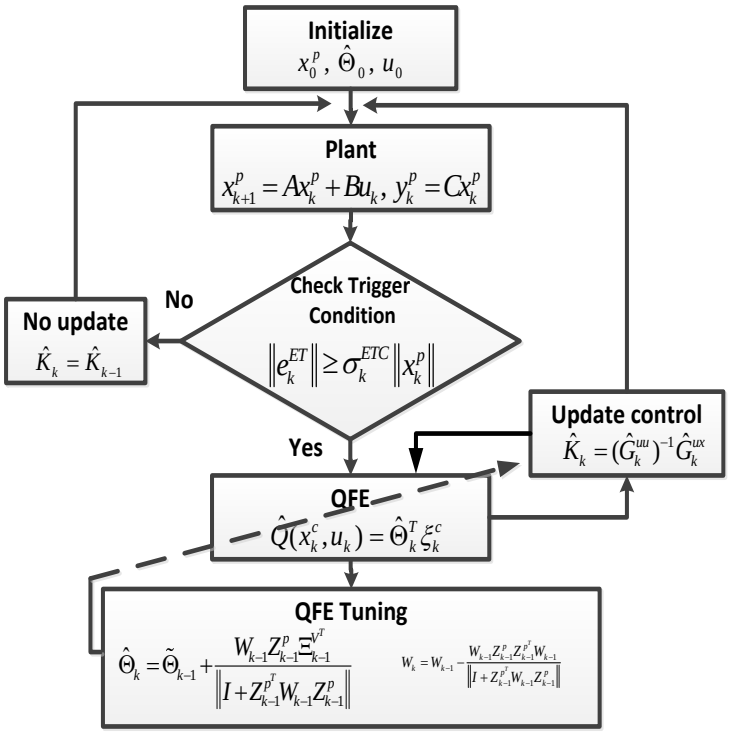

Figure 3. Implementation of event sampled Q-learning using state feedback.

5. EVENT SAMPLED OUTPUT FEEDBACK DESIGN

The placement of the observer, either at the plant or at the controller, plays a crucial role in the event sampled systems. An event sampled observer along with QFE at the controller will lead to a large number of event sampled instants for the observer state estimation error to converge to zero. From the computation and optimal performance point of view, we propose an adaptive observer at the sensor node. Since, the observer is connected to the sensor, it has access to the system output, y_k^p, at every time instant $k \in \mathbb{N}$. Therefore, the observer parameters are tuned at every time instant k unlike the aperiodic tuning in case of QFE. This further helps in faster convergence of the estimated observer state to the system state vector as mentioned in Section 3.

5.1 OBSERVER DESIGN AND PARAMTER CONVERGENCE

Consider the adaptive observer dynamics (9). The observer state estimation error dynamics by using (1) and (9) can be written as

$$e_{k+1}^x = (A - LC)e_k^x + \tilde{\psi}_k^T \phi_k, \tag{32}$$

where $\tilde{\psi}_k = [\tilde{A}_k \quad \tilde{B}_k \quad \tilde{L}_k]^T \in \Re^{q \times n}$ is the observer parameter estimation error with $\tilde{\psi}_k = \psi - \hat{\psi}_k$, $\tilde{A}_k = A - \hat{A}_k$, $\tilde{B}_k = B - \hat{B}_k$ and $\tilde{L}_k = L - \hat{L}_k$. The parameter matrix $\psi = [A \quad B \quad L]^T \in \Omega_\psi \subset \Re^{q \times n}$ is the ideal observer parameters where L is the ideal observer gain matrix. The observability of the system in Assumption 2.1 guarantees the existence of an ideal observer gain matrix L such that the matrix $A - LC$ is Schur. Further, it is assumed that the observer gain L satisfies $0 < l_o < \min\{\sqrt{1/2}, \sqrt{\alpha_\psi^o \bar{\phi}_{\min}^2 / (2\alpha_\psi^o \bar{\phi}_{\min}^2 + 1)}\} < 1$ where $l_0 = \|A - LC\|$, $c_{\min} = \sigma_{\min}(X^T C)$,

$\sigma_{\min}(\bullet)$ is the minimum singular value, $X \in \Re^{p_o \times n}$ is a constant matrix, and $0 < \bar{\phi}_{\min} \le \left\| \phi_k \phi_k^T \right\| / (1 + \phi_k^T \phi_k) \le 1$ ensured by the PE condition of the observer regression vector. Selection of these parameters ensures a faster convergence of the observer state error.

The observer output error dynamics from (32) is given by

$$e_{k+1}^y = C\left(A - LC\right)e_k^x + C\tilde{\psi}_k^T \phi_k . \tag{33}$$

A tuning law to tune the observer parameters at every sampling instant is selected as

$$\hat{\psi}_{k+1} = \hat{\psi}_k + (\alpha_\psi^o \phi_k e_{k+1}^{y^T} X / (1 + \phi_k^T \phi_k)), \tag{34}$$

where $\alpha_\psi^o > 0$ is the learning gain, and $X \in \Re^{p_o \times n}$ is a constant matrix to match the dimension and selected such that $\left\| X^T C \right\| \le 1$. The observer parameter estimation error dynamics can be computed from (34) as

$$\tilde{\psi}_{k+1} = \tilde{\psi}_k - (\alpha_\psi^o \phi_k e_{k+1}^{y^T} X / (1 + \phi_k^T \phi_k)). \tag{35}$$

Lemma 5.1. Consider the adaptive observer (9) in a parametric form (10) and let the Assumption 2.1 holds. Assume the initial observer parameters $\hat{\psi}_0$ are initialized in a compact set Ω_ψ. Suppose the observer parameters are updated by the tuning law (34). Then, the observer state estimation error e_k^x and the parameter estimation error $\tilde{\psi}_k$ converge asymptotically to zero provided the regression vector ϕ_k satisfies PE condition and the learning gain satisfies $0 < \alpha_\psi^o < \min\{c_{\min}/2,\ c_{\min}/2(1 + \| \phi_k \|^2)\}$.

Proof. Refer to the Appendix.

5.2 CONTROLLER DESIGN AND CLOSED-LOOP STABILITY

Similar to the state feedback case, the QFE parameters are tuned at the event sampled instants. Therefore, the observer based Bellman error (20) at event sampled instants with $x_k^{o,c} = x_k^o$ can be defined as $e_k^{o,V} = r(x_k^o, u_k) + \hat{\Theta}_k^T \Delta\xi_k^o,\ k = k_i$, where $r(x_k^o, u_k) = x_k^{o^T} P x_k^o + u_k^T R u_k$ and $\Delta\xi_k^o = \xi_{k+1}^o - \xi_k^o$. The augmented Bellman error can be defined as

$$\Xi_k^{o,V} = \Pi_k^o + \hat{\Theta}_k^T Z_k^o,\ \ k = k_i, \tag{36}$$

where $\Pi_k^o = [r(x_{k_i}^o, u_{k_i})\ \ r(x_{k_{i-1}}^o, u_{k_{i-1}})\ \ \cdots\ \ r(x_{k_{i-1-j}}^o, u_{k_{i-1-j}}]$ and $Z_k^o = [\Delta\xi_{k_i}^o\ \ \Delta\xi_{k_{i-1}}^o\ \ \cdots\ \ \Delta\xi_{k_{i-1-j}}^o]$ for $k = k_i$ with $0 < j < i$.

A tuning law to tune the observer based QFE parameter estimates at the event sampling instants is selected as

$$\hat{\Theta}_k = \begin{cases} \hat{\Theta}_{k-1} - (\alpha_V^o Z_{k-1}^o \Xi_{k-1}^{o,V^T} \big/ \left\| I + Z_{k-1}^{o^T} Z_{k-1}^o \right\|), & k = k_i, \\ \hat{\Theta}_{k-1}, & k_{i-1} < k < k_i. \end{cases} \tag{37}$$

where α_V^o is the learning gain. The QFE parameter estimation error dynamics by using (37) with a forwarded time step can be represented as

$$\tilde{\Theta}_{k+1} = \begin{cases} \tilde{\Theta}_k + (\alpha_V^o Z_k^o \Xi_k^{o,V^T} \big/ \left\| I + Z_k^{o^T} Z_k^o \right\|), & k = k_i, \\ \tilde{\Theta}_k, & k_i < k < k_{i+1}. \end{cases} \tag{38}$$

The observer based Bellman error ($e_k^{o,V}$) by using (36) and (6) at the event sampled instants $k = k_i$ is expressed in terms of the parameter estimation error $\tilde{\Theta}_k$ as $e_k^{o,V} = -\tilde{\Theta}_k^T \Delta\xi_k^o + \Theta^T (\Delta\xi_k^o - \Delta\xi_k^p) + f(x_k^o) - f(x_k^p)$ for $k = k_i$ where $f(x_k^o) = x_k^{o^T} P x_k^o$ and $f(x_k^p) = x_k^{p^T} P x_k^p$. The observer based augmented Bellman error (36) can be rewritten as

$$\Xi_k^{o,V} = -\tilde{\Theta}_k^T Z_k^o + \Theta^T (Z_k^o - Z_k^p) + F_k^o - F_k^p \,,\; k = k_i \,, \tag{39}$$

where $F_k^o = [x_{k_i}^{o^T} P x_{k_i}^o \quad x_{k_{i-1}}^{o^T} P x_{k_{i-1}}^o \quad \cdots \quad x_{k_{i-j-1}}^{o^T} P x_{k_{i-j-1}}^o]$ and $F_k^p = [x_{k_i}^{p^T} P x_{k_i}^p \quad x_{k_{i-1}}^{p^T} P x_{k_{i-1}}^p \quad \cdots \quad x_{k_{i-j-1}}^{p^T} P x_{k_{i-j-1}}^p]$.

The estimated control input (22) with the event-based observer state vector and the Q-function estimated parameters can be written as

$$u_k = -\hat{K}_k x_k^{o,c} = -(\hat{G}_k^{uu})^T \hat{G}_k^{ux} x_k^{o,c}, \;\; k_i \le k < k_{i+1}. \tag{40}$$

The closed-loop dynamics of the observer based system by using system dynamics (1) and control input (40), become

$$x_{k+1}^p = (A - B\hat{K}_k) x_k^p + B\hat{K}_k e_k^x + B\hat{K}_k e_k^{o,ET}, \;\; k \le k < k_{i+1}. \tag{41}$$

Consider the observer event sampling error (18). The triggering condition by using the observer state is selected as

$$\left\| e_k^{0,ET} \right\| \le \sigma_k^{o,ETC} \left\| x_k^o \right\|, \tag{42}$$

where $\sigma_k^{o,ETC} = \sqrt{\Gamma_{ET}^o (1 - 4\mu) \Big/ 4B_{\max}^2 \left\| \hat{K}_k \right\|^2}$ is the threshold coefficient, $0 < \Gamma_{ET}^o < 1$ and $\mu < 1/4$. Similar to the state feedback case, to ensure the estimated control gain $\left\| \hat{K}_k \right\|$ is nonzero the previous nonzero value is used to evaluate the threshold coefficient when the estimated gain $\left\| \hat{K}_k \right\|$ becomes zero. The event sampled instants are decided by the violation of the condition (42).

Theorem 5.2. Consider the uncertain LTI discrete-time system (1), the adaptive observer (9), and the observer based controller (40) represented as a closed-loop system (41). Let the Assumption 2.1 holds and the regression vectors ϕ_k and ξ_k^o satisfy the PE condition. Suppose $u_0 \in \Omega_u$ is the initial admissible control policy and the design parameters satisfy

$0 < \alpha_V^o < 1/3$ and $0 < \alpha_\psi^o < \min\{c_{\min}/2,\ c_{\min}/2(1+\|\phi_k\|^2)\}$. Let the state vector $x_k^{o,c}$, the QFE parameter vector $\hat{\Theta}_k$ are updated, respectively, by (17) and (37) at the violation of the triggering condition (42). Then, the closed-loop system state vector x_k^p and the observer state estimation error e_k^x, the QFE parameter estimation error $\tilde{\Theta}_k$, and the observer parameter estimation error $\tilde{\psi}_k$ converge to zero asymptotically for all $x_0^p \in \Omega_x$, $\hat{\Theta}_0 \in \Omega_\Theta$, and $\hat{\psi}_0 \in \Omega_\psi$ with event sampled instants $k_i \to \infty$ or, $k \to \infty$. Further, the estimated Q-function $\hat{Q}(x_k^p, u_k) \to Q^*(x_k^p, u_k^*)$ and estimated control input $u_k \to u_k^*$ as $k \to \infty$.

Proof. Refer to the Appendix.

In case of a discrete time system, the minimum inter-event time for both the state and the output feedback case is trivial and equal to the periodic sampling interval.

The event sampled Q-learning scheme designed in the previous section can be applied to the NCS in the presence of the networked induced time varying delays and random packet losses. The introduction of random parameters due to the communication network requires a stochastic analysis frame work for the event sampled Q-learning scheme. A complete design procedure along with simulation results are provided in the Appendix A of the dissertation. It was observed that the proposed event sampled stochastic Q-learning scheme shown a 30% reduction in computation and 56% reduction in network bandwidth usage.

6. SIMULATION RESULTS

The proposed optimal adaptive schemes are evaluated in this section by a numerical example. The benchmark example of the batch reactor is used here for simulation. The discrete-time version of the batch reactor with a sampling interval of $T_s = 0.01\text{sec}$ is given by

$$x_{k+1}^p = Ax_k^p + Bu_k, \quad y_k^p = Cx_k^p,$$

where $A = \begin{bmatrix} 1.0142 & -0.0018 & 0.0651 & -0.0546 \\ -0.0057 & 0.9582 & -0.0001 & 0.0067 \\ 0.0103 & 0.0417 & 0.9363 & 0.0563 \\ 0.0004 & 0.0417 & 0.0129 & 0.9797 \end{bmatrix}$, $B = \begin{bmatrix} 4.7798\times10^{-6} & -0.0010 \\ 0.0556 & 1.5316\times10^{-6} \\ 0.0125 & -0.0304 \\ 0.0125 & -0.0002 \end{bmatrix}$

and $C = \begin{bmatrix} 1 & 0 & 1 & -1 \\ 0 & 1 & 0 & 0 \end{bmatrix}$.

6.1 STATE FEEDBACK DESIGN

The state feedback design was evaluated first. A quadratic cost function was chosen as in (2) with the penalty matrices $P = I_{4\times4}$ and $R = I_{2\times2}$ where I denotes the identity matrix. The initial system states were selected as $x_0 = \begin{bmatrix} 0.1 & -0.1 & 0.3 & -0.5 \end{bmatrix}^T$. The initial parameter vector $\hat{\Theta}_0 \in \Re^{l_g=21}$ was chosen at random from a uniform distribution in the interval $[0,\ 1]$. The design parameters were $\beta = 2\times10^5$, $\mu = 0.3$, and $\Gamma = 0.1$. The PE condition was satisfied by adding a zero mean Gaussian noise with the control input. The simulation was conducted for 10 sec with a fixed sampling interval of 0.01 sec or 1000 sampling instants.

The event sampled optimal controller's performance is illustrated in Figure 4. The convergence of the system state and the threshold to zero are depicted in Figure 4 (a) and (b), respectively. The triggering condition is shown in Figure 4(b) evolved during the inter-event times and resets to zero at the event sampling instants. The cumulative number of event sampled instants plotted in Figure 4 (c) was found to be 108 out of 1000 sampling instants. This implies the mirror and controller QFE were updated only 108 times. Thus, the computation was reduced when compared to the traditional Q-learning based systems. Table 1 shows the comparison of computational load in terms of the additions and multiplications and a saving of 72% of the computation observed in the event sampled system when compared to its periodic counterpart.

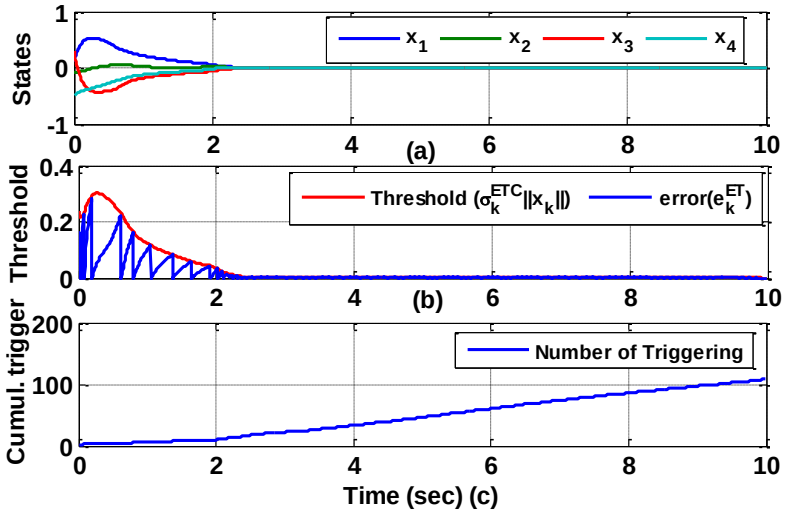

Figure 4. State feedback controller design: (a) convergence of state vector, (b) evolution of threshold and the event sampling error, (c) the total number of event sampled instants.

The control input is plotted in the Figure 5 (a). Figure 5 (b) shows the convergence of Bellman error to zero implies the Bellman equation is satisfied and the optimality is

achieved with aperiodic tuning of the QFE parameters. The convergence of the QFE parameter estimation error to zero is shown in Figure 5 (c).

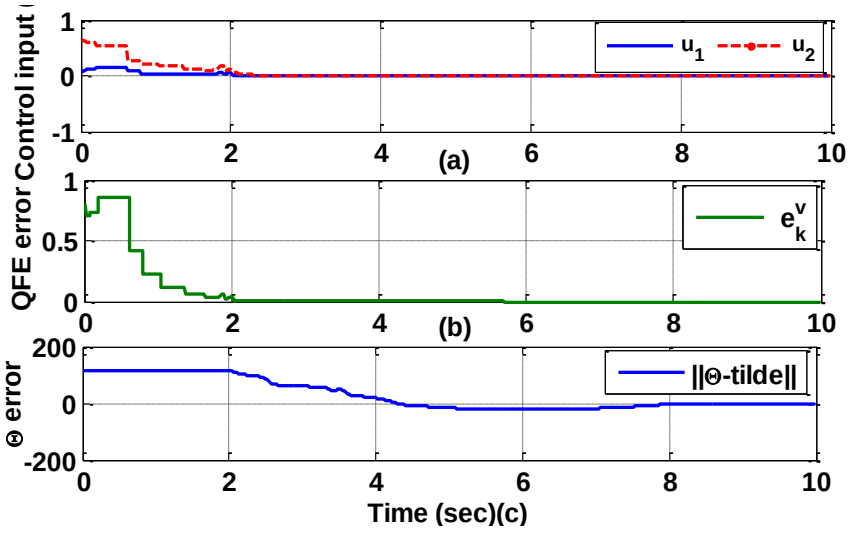

Figure 5. Optimal control input, (b) Bellman error, and (c) QFE parameter estimation error.

Table 1. Comparison of computational load between traditional periodic sampled and event sampled system

System		Traditional periodic sampled	Event-based non-periodic sampled
Samping instants		1000	108
Number of additions and Multiplications at every sampling instant	QFE	13	13
	Controller	3	3
	Update law(controller and trigger mechanism both)	65	65 x 2
	Trig. Con (periodic execution)	0	7
Total number of Computation		81000	22768

6.2 OUTPUT FEEDBACK DESIGN

The output feedback design was evaluated by selecting the following simulation parameters. The adaptive gains for the observer were $\alpha_{\Psi}^{0} = 0.01, \mu = 0.2$, and $\Gamma_{ET}^{o} = 0.1$.

The observer parameters were initialized at random from the uniform distribution in the interval $[0, 0.5]$. The initial observer states were given as $x_0^o = [.02, -.02, .03, -0.1]^T$. The remaining parameters for the state feedback were used here as well.

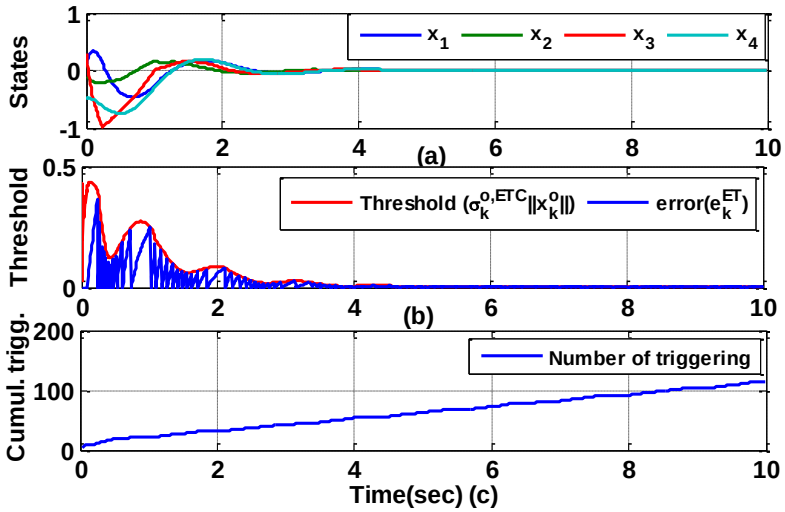

Figure 6. Output feedback controller design: (a) convergence of system states, (b) evolution of both the triggering condition threshold and the observer event sampling error, (c) the cumulative number of event sampled instants.

The observer-based output feedback controller performance is illustrated in Figures 6 and 7. The system states and the threshold are converged to zero as shown in Figure 6 (a) and (b). It was observed that the number of cumulative event sampled instants was increased to 115, as shown in Figure 6 (c), when compared to the state feedback case. This is due the additional uncertainty introduced by adaptive observer. The convergence of the observer based Bellman error and state estimation error are shown in Figure 7 (b) and (c).

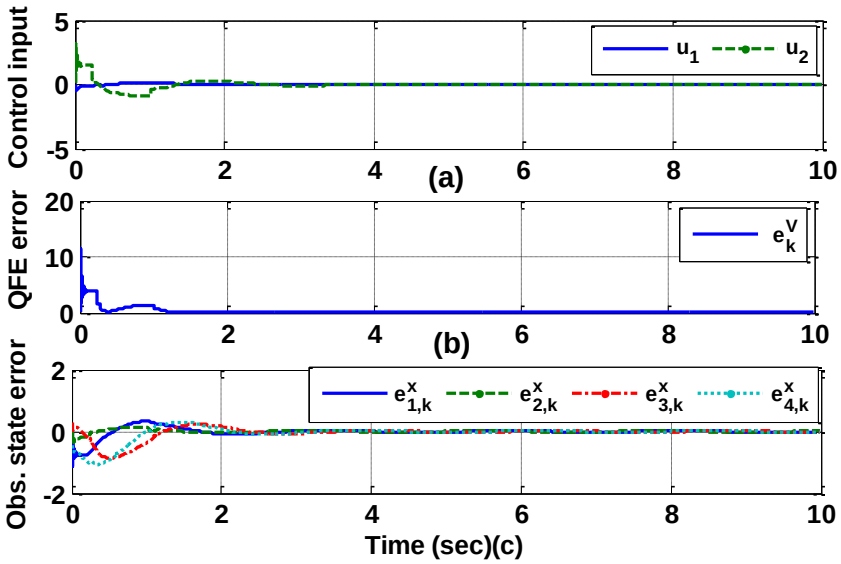

Figure 7. Evolution of (a) Optimal control input, (b) Bellman error, and (c) convergence of observer state error.

7. CONCLUSIONS

In this paper, two optimal control techniques with event sampled state and input vector for an uncertain linear discrete time system were presented. Both the state and the output feedback designs were able to regulate the system state vector without needing the system dynamics. The triggering conditions ensured sufficient number of events in both the cases for estimation of the QFE parameters. The aperiodic tuning guaranteed the convergence of the QFE parameter estimation errors as proved by the Lyapunov technique. The simulation results for both cases validated the analytical results by revealing the convergence of the closed-loop parameters and the reduction in computation. It was observed that the cumulative number of aperiodic sampled instants was dependent on the initial QFE parameters. In addition, the output feedback design triggered a more number of events than did the state feedback case.

8. REFERENCES

Al-Tamimi, A., Lewis, F. L., & Abu-Khalaf, M. (2007). Model-free Q-learning designs for linear discrete-time zero-sum games with application to H-infinity control. Automatica, 43, 473–481.

Barto, A. G., Sutton, R. S., & Anderson, C. W. (1983). Neuronlike elements that can solve difficult learning control problems. *IEEE Transactions on Systems Man and Cybernetics, SMC-B*, 835–846.

Beard, R. (1995). Improving the closed-loop performance of nonlinear systems. Ph.D. dissertation, RPI, USA.

Bertsekas, D. P., & Tsitsiklis, J. N. (1996). Neuro-dynamic programming. MA: Athena Scientific.

Bradtke, S.J., Ydestie, B.E., & Barto, A.G. (1994). Adaptive linear quadratic control using policy iteration. In *Proceedings of the American Control Conference,* (pp. 3475–3476).

Chatterjee, D. & Liberzon, D. (2006). Stability analysis of deterministic and stochastic switched systems via a comparison principle and multiple Lyapunov functions. *SIAM Journal on Control and Optimization,* 45 (1), 174-206.

Cogill, R. (2009). Event-based control using quadratic approximate value functions. In *Proceedings of the Joint IEEE Conference on Decision and Control and Chinese Control Conference,* (pp. 5883–5888).

Dierks, T., & Jagannathan, S. (2012). Online optimal control of affine nonlinear discrete-time system with unknown internal dynamics by using time-based policy update. *IEEE Transaction on Neural Network and Learning System,* 23 (7), 1118-1129.

Donkers, M. C. F., & Heemels, W. P. M. H. (2012). Output-based event-triggered control with guaranteed L∞-gain and improved and decentralized event-triggering. *IEEE Transactions on Automatic Control*, 57 (6), 1362–1376.

Goodwin, G.C., &Sin, K.S. (1984) *Adaptive filtering prediction and control*. 1st edition. New York: Dover publications.

Hagen, S., & Krose, B. (1998). Linear quadratic regulation using reinforcement learning. In *Proceedings of the Belgian-Dutch conference on mechanical learning* (pp. 39–46).

Imer, O.C., & Basar, T. (2006). To measure or to control: optimal control with scheduled measurements and controls. In *Proceedings of the American Control Conference,* (pp. 14-16).

Jagannathan, S. (2006). Neural Network Control of Nonlinear Discrete-time Systems. Boca Raton, FL: CRC Press.

Lewis, F. L. & Syrmos, V. L. (1995). Optimal Control, 2nd edition. New York: Wiley.

Molin, A., & Hirche, S. (2013). On the optimality of cerntainty equivalence for event-triggered control systems. *IEEE Transaction on Automatic Control* 58(2), 470-474.

Sahoo, A., and Jagannathan, S., (2014). Event-triggered optimal regulation of uncertain linear discrete-time systems by using Q-learning scheme. In *Proceedings of the IEEE Conference on Decision and Control* (pp. 1233-1238).

Si, J., Barto, A., Powell, W., & Wunsch, D. (2004). Handbook of learning and approximate dynamic programming. New Jersey: Wiley.

Tabuada, P. (2007). Event-triggered real-time scheduling of stabilizing control tasks. *IEEE Transactions on Automatic Control*, 52(9), 1680–1685.

Wang, Fei-Yue, Jin, N., Liu, D., & Wei, Q. (2011). Adaptive dynamic programming for finite-horizon optimal control of discrete-time nonlinear systems with ε -error bound. *IEEE Transaction on Neural Networks*, 22 (1), 24–36.

Wang, X., & Lemmon, M. D., (2011). On event design in event-triggered feedback systems. *Automatica*, 47(10), 2319-2322.

Watkins, C. J. C. H. (1989). Learning from delayed rewards. PhD thesis, Cambridge University, Cam- Bridge, England.

Werbos, P. J. (1992). Approximate dynamic programming for real-time control and neural modeling. In: D. A. White, & D. A. Sofge (Eds.), Handbook of intelligent control New York: Van Nostrand Reinhold.

Xu, Hao, Jagannathan, S., & Lewis, F. L. (2012). Stochastic optimal control of unknown linear networked control system in the presence of random delays and packet losses. *Automatica*, 48(6), 1017-1030.

Zhao, Q., Xu, Hao, & Jagannathan, S., (2014). Finite-horizon near optimal control of uncertain linear discrete-time system. *Optimal Control Applications and Methods*, DOI: 10.1002/oca.2143.

APPENDIX

Proof of Lemma 4.4. The proof is carried out by considering both the cases of triggering condition, i.e., at the event sampled instants ($k = k_i$) and inter-event times ($k_i < k < k_{i+1}$), because of the aperiodic tuning of the QFE parameters. A single Lyapunov function is used to evaluate both the cases and the asymptotic stability is shown by combining at the end.

Consider the Lyapunov function candidate given as

$$L_{\tilde{\Theta},k} = \tilde{\Theta}_k^T W_k^{-1} \tilde{\Theta}_k . \tag{A.1}$$

where W_k is a positive definite matrix as defined in (25).

Case I: At the event sampled instants ($k = k_i$)

In this case, the QFE parameters are tuned by using (24) and (25) for the case $k = k_i$. The QFE parameter estimation error dynamics (26) with the augmented Bellman error (27) , can be written as

$$\tilde{\Theta}_{k+1} = \left(I - \frac{W_k Z_k^p Z_k^{p^T}}{\left\| I + Z_k^{p^T} W_k Z_k^p \right\|} \right) \tilde{\Theta}_k . \tag{A.2}$$

Again, Eq. (25) can be expressed as

$$W_{k+1} W_k^{-1} = I - \frac{W_k Z_k^p Z_k^{p^T}}{\left\| I + Z_k^{p^T} W_k Z_k^p \right\|} . \tag{A.3}$$

Substituting (A.3) in (A.2), the QFE parameter estimation error dynamics become

$$\tilde{\Theta}_{k+1} = W_{k+1} W_k^{-1} \tilde{\Theta}_k . \tag{A.4}$$

Consider the Lyapunov function (A.1). The first difference $\Delta L_{\tilde{\Theta},k}$ along (A.4) can be expressed as

$$\Delta L_{\tilde{\Theta},k} = \tilde{\Theta}_{k+1}^T W_{k+1}^{-1} W_{k+1} W_k^{-1} \tilde{\Theta}_k - \tilde{\Theta}_k^T W_k^{-1} \tilde{\Theta}_k = (\tilde{\Theta}_{k+1} - \tilde{\Theta}_k)^T W_k^{-1} \tilde{\Theta}_k.$$

Substituting (A.2) in the above equation one can arrive at

$$\Delta L_{\tilde{\Theta},k} = \left[\left(I - \frac{W_k Z_k^p Z_k^{p^T}}{\left\| I + Z_k^{p^T} W_k Z_k^p \right\|} \right) \tilde{\Theta}_k - \tilde{\Theta}_k \right]^T W_k^{-1} \tilde{\Theta}_k$$

$$= - \frac{\tilde{\Theta}_k^T Z_k^p Z_k^{p^T} \tilde{\Theta}_k}{\left\| I + Z_k^{p^T} W_k Z_k^p \right\|}. \tag{A.5}$$

The regression vector Z_k^p satisfies the PE condition as discussed in Remark 4.2. Therefore, it holds that $0 < \bar{Z}_{\min} \le \frac{\left\| Z_k^p Z_k^{p^T} \right\|}{\left\| I + Z_k^{p^T} W_k Z_k^p \right\|} < 1$. The first difference in (A.5) by using the above observation is upper bounded as

$$\Delta L_{\tilde{\Theta},k} \le - \bar{Z}_{\min}^2 \left\| \tilde{\Theta}_k \right\|^2 < 0. \tag{A.6}$$

This implies, the positive definite Lyapunov function $L_{\tilde{\Theta},k}$ is a non-increasing function i.e., $L_{\tilde{\Theta},k_i+1} < L_{\tilde{\Theta},k_i}$.

Case II: During the inter-event times ($k_i < k < k_{i+1}$)

In this case, the QFE parameters are not tuned and held at their previous values. Consider the same Lyapunov function in (A.1). The first difference along (25) and (26) for $k_i < k < k_{i+1}$ is given by

$$\Delta L_{\tilde{\Theta},k} = \tilde{\Theta}_{k+1}^T W_{k+1}^{-1} \tilde{\Theta}_{k+1} - \tilde{\Theta}_k^T W_k^{-1} \tilde{\Theta}_k = 0, \; k_i < k < k_{i+1}. \tag{A.7}$$

By Lyapunov theorem (Jagannathan, 2006), the QFE parameter estimation error $\tilde{\Theta}_k$ remains constant during the inter-event times. Now by combining both the cases for

the interval $k_i \le k < k_{i+1}$, we will show that the QFE parameter estimation error converge to zero asymptotically.

From Case I, the first difference (A.6) at the event sampled instants $k = k_i$ can be expressed as

$$\Delta L_{\tilde{\Theta},k_i} = \tilde{\Theta}_{k_i+1}^T W_{k_i+1}^{-1} \tilde{\Theta}_{k_i+1} - \tilde{\Theta}_{k_i}^T W_{k_i}^{-1} \tilde{\Theta}_{k_i} \ge \lambda_{\min}(W_{k_i+1}^{-1}) \left\| \tilde{\Theta}_{k_i+1} \right\|^2 - \lambda_{\max}(W_{k_i}^{-1}) \left\| \tilde{\Theta}_{k_i} \right\|^2. \quad \text{(A.8)}$$

By using matrix inversion Lemma $(\mathrm{A} - BD^{-1}C)^{-1} = A^{-1} + A^{-1}B(D - CA^{-1}B)^{-1}CA$ (Goodwin, & Sin, 1984), the inverse of (25) for $k = k_i$ can be expressed as

$$W_{k_i+1}^{-1} = W_{k_i}^{-1} + \Delta Z_{k_i}^p (\| I + Z_{k_i}^{p^T} W_{k_i} Z_{k_i}^p \| I - Z_{k_i}^{p^T} W_{k_i} Z_{k_i}^p)^{-1} \Delta Z_{k_i}^{p^T},$$

where $(\| I + Z_{k_i}^{p^T} W_{k_i} Z_{k_i}^p \| I - Z_{k_i}^{p^T} W_{k_i} Z_{k_i}^p)$ is positive definite matrix. Therefore, it holds that $\lambda_{\min}(W_{k_i+1}^{-1}) \ge \lambda_{\min}(W_{k_i}^{-1}) \ge \lambda_{\min}(W_0^{-1})$ and $\lambda_{\max}(W_{k_i+1}^{-1}) \ge \lambda_{\max}(W_{k_i}^{-1}) \ge \lambda_{\max}(W_0^{-1})$. By using the above relations, (A.8) satisfies

$$\Delta L_{\tilde{\Theta},k_i} \ge \lambda_{\min}(W_{k_i+1}^{-1}) \left\| \tilde{\Theta}_{k_i+1} \right\|^2 - \lambda_{\max}(W_{k_i}^{-1}) \left\| \tilde{\Theta}_{k_i} \right\|^2 \ge \lambda_{\min}(W_0^{-1}) \left\| \tilde{\Theta}_{k_i+1} \right\|^2 - \lambda_{\max}(W_0^{-1}) \left\| \tilde{\Theta}_{k_i} \right\|^2. \quad \text{(A.9)}$$

Now, comparing (A.6) with(A.9), it holds that

$$\lambda_{\min}(W_0^{-1}) \left\| \tilde{\Theta}_{k_i+1} \right\|^2 - \lambda_{\max}(W_0^{-1}) \left\| \tilde{\Theta}_{k_i} \right\|^2 \le -\bar{Z}_{\min}^2 \left\| \tilde{\Theta}_{k_i} \right\|^2. \quad \text{(A.10)}$$

From (25), $W_0 = \beta I$ implies $\lambda_{\min}(W_0^{-1}) = \lambda_{\max}(W_0^{-1}) = \beta$. Then, (A.10) becomes

$$\left\| \tilde{\Theta}_{k_i+1} \right\|^2 \le (1 - (\bar{Z}_{\min}^2 / \beta)) \left\| \tilde{\Theta}_{k_i} \right\|^2. \quad \text{(A.11)}$$

Recalling Case II, $\tilde{\Theta}_k$ remains constant during the inter-event time and from (26), we have $\tilde{\Theta}_{k_i} = \tilde{\Theta}_{k_{i-1}+1}$, $k_{i-1} < k < k_i$. Therefore, (A.11) can be rewritten as

$$\left\| \tilde{\Theta}_{k_i+1} \right\|^2 \le \left((1 - (\bar{Z}_{\min}^2 / \beta)) \right) \left\| \tilde{\Theta}_{k_{i-1}+1} \right\|^2.$$

Solving the difference inequality with initial value $\left\|\tilde{\Theta}_{k_0}\right\| = \left\|\tilde{\Theta}_0\right\| = B_{\tilde{\Theta}0}$ and by comparison lemma (Chatterjee & Liberzon, 2006)

$$\left\|\tilde{\Theta}_{k_i+1}\right\|^2 \leq \alpha^{i+1} B_{\tilde{\Theta}0}^2 \equiv B_{\tilde{\Theta},k_i+1}, \tag{A.12}$$

where $\alpha = 1-(\bar{Z}_{\min}^2/\beta)$ and $B_{\tilde{\Theta},k_i+1}$ is a piecewise constant function and remains constant for i^{th} inter-event time. Since, $0<\alpha<1$, $B_{\tilde{\Theta},k_i+1}$ is converging sequence, i.e., $B_{\tilde{\Theta},k_{i+1}+1} < B_{\tilde{\Theta},k_i+1}$, $i=1,2,\cdots$ and $B_{\tilde{\Theta},k_i+1} \to 0$ as $i \to \infty$. Therefore, $\tilde{\Theta}_k \to 0$ with $i \to \infty$ or $k_i \to \infty$. Since k_i is a subsequence of k , the QFE parameter estimation error $\tilde{\Theta}_k \to 0$ as $k \to \infty$. ■

Proof of Theorem 4.7. To show the asymptotic stability of the closed-loop system we will evaluate a Lyapunov function for both cases of the triggering condition and will show that (31) holds. Consider the Lyapunov function candidate

$$L^s{}_k = \Lambda_1 L_{x^p,k} + \Lambda_2 L_{\tilde{\Theta},k}, \tag{A.13}$$

where $L_{x^p,k} = x_k^{p^T} x_k^p$ and $L_{\tilde{\Theta},k} = \tilde{\Theta}_k^T W_k^{-1} \tilde{\Theta}_k$ with $\Lambda_1 = \pi/2B_{\max}^2 l_{\tilde{\Theta}}$ and $\Lambda_2 = 2\pi$ where $0<\pi<1$ and $l_{\tilde{\Theta}}$ is a positive constants .

Case 1. At the sampled instants ($k = k_i$)

We will evaluate each term in the Lyapunov function candidate (A.13) individually and combine them to compute the overall first difference, for simplicity.

At the event sampled instants with $e_k^{ET} = 0$, the closed loop system dynamics (28) can be expressed as

$$x_{k+1}^p = Ax_k^p - B\hat{K}_k x_k^p, \quad k = k_i. \tag{A.14}$$

Consider the first term in (A.13). The first difference, $\Delta L_{x^p,k} = x_{k+1}^{p^T} x_{k+1}^p - x_k^{p^T} x_k^p$, along the system dynamics (A.14) with the relation $\tilde{K}_k = K^* - \hat{K}_k$ and Cauchy-Schwartz (C-S) inequality , can be expressed as

$$\Delta L_{x^p,k} \le 2x_k^{p^T}(A-BK^*)^T(A-BK^*)x_k^p + 2x_k^{p^T}(B\tilde{K}_k)^T(B\tilde{K}_k)x_k^p - x_k^{p^T}x_k^p.$$

With simple mathematical operations using Lemma 4.6 and norm, one can arrive at

$$\Delta L_{x^p,k} \le -(1-2\mu)\left\|x_k^p\right\|^2 + 2B_{\max}^2\left\|\tilde{K}_k x_k^p\right\|^2. \tag{A.15}$$

Next, considering the second term, $L_{\tilde{\Theta},k}$, in (A.13), the first difference is same as in (A.5) of Lemma 4.4.

At the final step, combining the individual first differences (A.5) and (A.14), the overall first difference $\Delta L^s{}_k = \Lambda_1 \Delta L_{x^p,k} + \Lambda_2 \Delta L_{\tilde{\Theta},k}$ becomes

$$\Delta L^s{}_k \le -(1-2\mu)\left\|x_k^p\right\|^2 + 2B_{\max}^2\left\|\tilde{K}_k x_k^p\right\|^2 - \tilde{\Theta}_k^T Z_k^p Z_k^{pT}\tilde{\Theta}_k \Big/ \left\|I + Z_k^{p^T} W_k Z_k^p\right\|$$

Since the initial control input is admissible, $\tilde{K}_k$ is a function of $\tilde{\Theta}_k$ and Z_k^p is a function of x_k^p, by Lipschitz continuity $\left\|\tilde{K}_k x_k^p\right\|^2 \le l_{\tilde{\Theta}}\left\|\tilde{\Theta}_k^T Z_k^p Z_k^{pT}\tilde{\Theta}_k\right\| \Big/ \left\|I + Z_k^{p^T} W_k Z_k^p\right\|$ holds where $l_{\tilde{\Theta}} > 0$ is a positive constant. With the above facts and recalling the definition Λ_1 and Λ_2 in (A.13), the overall first difference can be written as

$$\Delta L_k^s \le -(1-2\mu)\Lambda_1\left\|x_k^p\right\|^2 - \pi\bar{Z}_{\min}^2\left\|\tilde{\Theta}_k\right\|^2 < 0, \tag{A.16}$$

where $0 < \pi < 1$, $0 < \mu < 1/2$ and $\bar{Z}_{\min}$ is defined in (A.5). By Lyapunov theorem (Jagannathan, 2006), the Lyapunov function is a non-increasing function, i.e., $L_{k_i+1}^s < L_{k_i}^s$.

Case 2. During the inter-event times ($k_i < k < k_{i+1}$)

Consider the same Lyapunov function (A.13) as in Case 1. The system dynamics during the inter-event times become

$$x_{k+1}^p = Ax_k^p - B\hat{K}_k x_k^p + B\hat{K}_k e_k^{ET},\ k_i < k < k_{i+1}. \tag{A.17}$$

The first difference of the first term, with system dynamics (A.17), Lemma 4.6 and C-S inequality can be expressed as

$$\Delta L_{x^p,k} \le -(1-3\mu)\left\|x_k^p\right\|^2 + 3B_{\max}^2\left\|\tilde{K}_k x_k^p\right\|^2 + 3B_{\max}^2\left\|\hat{K}_k\right\|^2\left\|e_k^{ET}\right\|^2.$$

Recalling the triggering condition (29) one can reach at

$$\Delta L_{x^p,k} \le -(1-\Gamma)(1-3\mu)\left\|x_k^p\right\|^2 + 3B_{\max}^2 l_{\tilde{\Theta}}\frac{\left\|\tilde{\Theta}_k^T Z_k^p Z_k^{pT}\tilde{\Theta}_k\right\|}{\left\|I + Z_k^{pT} W_k Z_k^p\right\|}. \tag{A.18}$$

Moving on for the second term, the first difference along (26) for $k_i < k < k_{i+1}$ remains same as in (A.7). Combing the individual first differences (A.7) and (A.18), the overall first difference $\Delta L^s{}_k = \Lambda_1 \Delta L_{x^p,k} + \Lambda_2 \Delta L_{\tilde{\Theta},k}$ is expressed as

$$\Delta L_k^s \le -(1-\Gamma)(1-3\mu)\Lambda_1\left\|x_k^p\right\|^2 + 3\Lambda_1 B_{\max}^2 l_{\tilde{\Theta}}\left\|\tilde{\Theta}_k\right\|^2. \tag{A.19}$$

From (A.12) in Lemma 4.4, $\tilde{\Theta}_k$ remains constant for $k_i < k < k_{i+1}$. Thus, $\left\|\tilde{\Theta}_k\right\|^2 = \left\|\tilde{\Theta}_{k_i+1}\right\|^2 \le B_{\tilde{\Theta},k_i+1}$ for $k_i < k < k_{i+1}$. Substituting the inequality in (A.19), the first difference

$$\Delta L_k^s \le -(1-\Gamma)(1-3\mu)\Lambda_1\left\|x_k^p\right\|^2 + B_{\tilde{\Theta},k_i+1}^s, \tag{A.20}$$

where $B_{\tilde{\Theta},k_i+1}^s = 3\Lambda_1 B_{\max}^2 l_{\tilde{\Theta}} B_{\tilde{\Theta},k_i+1}$. From (A.20), the first difference of the Lyapunov function $\Delta L_k^s < 0$, as long as

$$\left\|x_k^p\right\| > \sqrt{B^s_{\tilde{\Theta},k_i+1} \big/ (1-\Gamma)(1-3\mu)\Lambda_1} = B^c_{x^p,k_i+1}.$$

By Lyapunov theorem (Jagannathan, 2006), the system state, x_k^p and QFE parameter error $\tilde{\Theta}_k$ are bounded. Further, system state x_k^p converges to the ball of radius $B^c_{x^p,k_i+1}$ in a finite time and $\tilde{\Theta}_k$ remains constant.

The bound for the Lyapunov function (A.13) for $k_i < k < k_{i+1}$ can obtained by using the bounds for x_k^p and $\tilde{\Theta}_k$ computed in (A.20) and (A.12), respectively . It is given by

$$B_{L,k} = \Lambda_1 (B^c_{x^p,k_i+1})^2 + \Lambda_2 (B_{\tilde{\Theta},k_i+1})^2 \text{ for } k_i < k < k_{i+1}. \tag{A.21}$$

It follows that the Lyapunov function L_k^s for $k_i < k < k_{i+1}$ converges to the bound $B_{L,k}$ in a finite time and stay within $B_{L,k}$.

Now, from Case I and Case II, we will show the existence of a function $h(k)$ such that (31) holds to prove the asymptotic convergence of x_k^p and $\tilde{\Theta}_k$. With this effect, define a piecewise continuous function

$$h(k) = \max\{L_k^s, B_{L,k}\},\ k \in \mathbb{N}. \tag{A.22}$$

It is clear that $h(k) \geq L_k^s$ for all $k \in \mathbb{N}$. From Lemma 4.4, $B_{\tilde{\Theta},k_i+1} \to 0$ with event sampled instants $k_i \to \infty$. Using this relation and (A.20) in Case II , we have $B^s_{\tilde{\Theta},k_i+1} \to 0$ and, hence, $B^c_{x^p,k_i+1} \to 0$ as $k_i \to \infty$. Therefore, it follows form (A.21) that the bound $B_{L,k} \to 0$ as $k_i \to \infty$. Since, the Lyapunov function $L^s_{k_i+1} < L^s_{k_i}$ for $k = k_i$ and $L_k^s \to B_{L,k}$,

$k_i < k < k_{i+1}$, $L_k^s \to 0$ as $k_i \to \infty$. Consequently, the upper bound functions $h(k) \to 0$ as $k_i \to \infty$. Since k_i is a subsequence of $k \in \mathbb{N}$, by extension $h(k) \to 0$ as $k \to \infty$.

Finally, the convergence of $\hat{Q}(x_k^p, u_k) \to Q^*(x_k^p, u_k^*)$ can be seen by considering

$$\begin{aligned}\left\|Q^* - \hat{Q}\right\| &= \left\|\Theta \xi_k^p - \hat{\Theta}_k \xi_k^c\right\| = \left\|\tilde{\Theta}_k \xi_k^p + \hat{\Theta}_k (\xi_k^p - \xi_k^c)\right\| \\ &\le \left\|\tilde{\Theta}_k\right\| \left\|\xi_k^p\right\| + L_\xi \left\|\hat{\Theta}_k\right\| \left\|e_k^{ET}\right\| \le \left\|\tilde{\Theta}_k\right\| \left\|\xi_k^p\right\| + L_\xi \sigma_k^{ETC} \left\|\hat{\Theta}_k\right\| \left\|x_k^p\right\|,\end{aligned}$$

where L_ξ is a positive constant. Since, $x_k^p \to 0$ and $\tilde{\Theta}_k \to 0$ as $k \to \infty$, imply $\hat{Q}(x_k^p, u_k) \to Q^*(x_k^p, u_k^*)$. Similarly, to show $u_k \to u_k^*$, consider the difference

$$\left\|u_k^* - u_k\right\| = \left\|K^* x_k^c - \hat{K}_k x_k^c\right\| = \left\|\tilde{K}_k x_k^c\right\| \le C_{\tilde{\Theta}} \left\|\tilde{\Theta}_k\right\| \left\|x_k^c\right\|.$$

with $C_{\tilde{\Theta}} > 0$ is a constant. Since $\tilde{\Theta}_k \to 0$ as $k \to \infty$, implies $\left\|u_k^* - u_k\right\| \to 0$ as $k \to \infty$,or $u_k \to u_k^*$ as $k \to \infty$. ■

Proof of Lemma 5.1. Consider the Lyapunov function candidate given by,

$$L_{O,k}^o = \Lambda_3 L_{e^x,k}^o + L_{\tilde{\psi},k}^o \ , \tag{A.23}$$

where $L_{e^x,k}^o = e_k^{x^T} e_k^x$ and $L_{\tilde{\psi},k}^o = tr\{\tilde{\psi}_k^T \tilde{\psi}_k\}$. The positive constant $\Lambda_3 = \alpha_\psi^{o^2} \bar{\phi}_{\min}$ where $0 < \bar{\phi}_{\min} \le \left\|\phi_k \phi_k^T\right\| / (1 + \phi_k^T \phi_k) \le 1$ is ensured by the PE condition of the regression vector as discussed in Remark 4.2.

Consider the first term, $L_{e^x,k}^o = e_k^{x^T} e_k^x$, of the Lyapunov function (A.23). The first difference, $\Delta L_{e^x,k}^o = e_{k+1}^{x^T} e_{k+1}^x - e_k^{x^T} e_k^x$, along the state estimation error dynamics (32) becomes

$$\Delta L_{e^x,k}^o = \left((A - LC) e_k^x + \tilde{\psi}_k^T \phi_k\right)^T \left((A - LC) e_k^x + \tilde{\psi}_k^T \phi_k\right) - e_k^{x^T} e_k^x.$$

After simple mathematical manipulation with C-S inequality and norm, the first difference can be written as

$$\Delta L^o_{e^x,k} \le -(1-2l_0^2)\left\|e^x_k\right\|^2 + 2\left\|\phi_k\right\|^2\left\|\tilde{\psi}_k\right\|^2. \tag{A.24}$$

where $l_0 = \|A-LC\|$ and $1-2l_0^2 > 0$ by selecting $0 < l_o < \sqrt{1/2}$.

Considering the second term, $L^o_{\tilde{\psi},k} = tr\{\tilde{\psi}^T_k\tilde{\psi}_k\}$, the first difference along the dynamics of the observer parameter estimation error (35) becomes

$$\Delta L^o_{\tilde{\psi},k} = tr\left\{(\tilde{\psi}_k - \alpha^o_\psi\phi_k e^{y^T}_{k+1}X\big/(1+\phi^T_k\phi_k))^T(\tilde{\psi}_k - (\alpha^o_\psi\phi_k\, e^{y^T}_{k+1}X)\big/(1+\phi^T_k\phi_k))\right\} - tr\left\{\tilde{\psi}^T_k\tilde{\psi}_k\right\}.$$

After simple mathematical operations using the output error dynamics (33), C-S inequality and the fact $\left\|\phi_k\right\|^2\big/(1+\left\|\phi_k\right\|^2) \le 1$, the first difference is upper bounded as

$$\begin{aligned}\Delta L^o_{\tilde{\psi},k} \le\ & -c_{\min}\alpha^o_\psi tr\left\{\tilde{\psi}^T_k\phi_k\phi^T_k\tilde{\psi}_k\big/(1+\phi^T_k\phi_k)\right\} + 2\alpha^{o^2}_\psi tr\left\{\tilde{\psi}^T_k\phi_k\phi^T_k\tilde{\psi}_k\big/(1+\phi^T_k\phi_k)\right\}\\ & +2\alpha^{o^2}_\psi tr\left\{(X^TC(A-LC)e^x_k)(X^TC(A-LC)e^x_k)^T\right\}\\ & +\alpha^o_\psi tr\left\{((A-LC)e^x_k)((A-LC)e^x_k)^TC^TX\right\}\\ \le\ & -\alpha^o_\psi\bar{\phi}^2_{\min}(c_{\min}-2\alpha^o_\psi)\left\|\tilde{\psi}_k\right\|^2 + \alpha^o_\psi l_0^2(1+2\alpha^o_\psi)\left\|e^x_k\right\|^2.\end{aligned} \tag{A.25}$$

Eq. (A.25) is reached by using the definition $\bar{\phi}_{\min}$ from (A.23).

The overall first difference $\Delta L^o_{O,k} = \Lambda_3\Delta L^o_{e^x,k} + \Delta L^o_{\tilde{\psi},k}$, from (A.24), (A.25) becomes

$$\Delta L^o_{O,k} \le -\Lambda_3(1-2l_o^{\,2})\left\|e^x_k\right\|^2 + 2\Lambda_3\left\|\phi_k\right\|^2\left\|\tilde{\psi}_k\right\|^2 - \alpha^o_\psi\bar{\phi}^2_{\min}(c_{\min}-2\alpha^o_\psi)\left\|\tilde{\psi}_k\right\|^2 + \alpha^o_\psi l_0^2(1+2\alpha^o_\psi)\left\|e^x_k\right\|^2.$$

Substituting Λ_3 from (A.23), the overall first difference

$$\Delta L^o_{O,k} \le -\varpi_1\left\|e^x_k\right\|^2 - \varpi_2\left\|\tilde{\psi}_k\right\|^2. \tag{A.26}$$

where $\varpi_1 = \left(1-2l_o^{\,2} - (l_0^2/\alpha^o_\psi\bar{\phi}^2_{\min})\right)$ and $\varpi_2 = \alpha^o_\psi\bar{\phi}^2_{\min}(c_{\min}-2\alpha^o_\psi-2\alpha^o_\psi\left\|\phi_k\right\|^2)$ for brevity. Note that, $\varpi_1 > 0$ and $\varpi_2 > 0$ by the choice of α^o_ψ and l_o defined earlier.

The overall first difference (A.26) of the Lyapunov function (A.23) is less than zero, i.e., $\Delta L_{O,k} < 0$. Therefore, by Lyapunov theorem (Jagannathan, 2006), $e_k^x \to 0$ and $\tilde{\psi}_k \to 0$ as $k \to \infty$. ■

Proof of Theorem 5.2. Similar to the state feedback case, we will consider both the cases for the triggering condition to show that (31) holds by defining an upper bound function.

Case 1. At the event sampled instants ($k = k_i$)

Consider the Lyapunov function candidate given by

$$L_k^o = L_{x^p,k}^o + \Sigma_{o1} L_{O,k}^o + \Sigma_{\Theta 1} L_{\tilde{\Theta}1,k}^o + \Sigma_{o2} L_{O2,k}^o + \Sigma_{\tilde{\Theta}2} L_{\tilde{\Theta}_2,k}^o, \tag{A.27}$$

where $L_{x^p,k}^o = x_k^{pT} x_k^p$, $L_{\tilde{\Theta}1,k}^o = \tilde{\Theta}_k^T \tilde{\Theta}_k$, $L_{\tilde{\Theta}2,k}^o = (\tilde{\Theta}_k^T \tilde{\Theta}_k)^2$, $L_{O2,k}^o = (L_{O,k}^o)^2$, and $L_{O,k}^o$ is given in (A.23). The constant coefficients are defined as

$$\begin{aligned}
\Sigma_{o1} &= 2\left(6B_{\max}^2 K_M^2 + \Sigma_{\Theta 1}\varpi_4\right)\Big/\Lambda_3\varpi_1, \\
\Sigma_{\Theta 1} &= 4B_{\max}^2 l_{\tilde{\Theta}_o}\Big/\varpi_3,, \\
\Sigma_{o2} &= \left(6B_{\max}^2\varpi_3 + \Sigma_{\tilde{\Theta}2}\varpi_4{}^2\left(2-\varpi_3\right)\right)\Big/2\varpi_3\Lambda_3^2\,\varpi_1\left(1-\varpi_1\right), \\
\Sigma_{\tilde{\Theta}2} &= 8B_{\max}^2 l_{\tilde{\Theta}_o}^2\Big/\varpi_3(2-\varpi_3)
\end{aligned}$$

where $\varpi_1 > 0$ and $\varpi_2 > 0$ as defined in (A.26) and $\varpi_3 = \alpha_V^o(1-3\alpha_V^o)\bar{Z}_{\min}^{o^2}$ and $\varpi_4 = \alpha_V^o(2+3\alpha_V^o)(L_F^2 + \Theta_M^2 L_Z^2)$.

Considering the first term, $L_{x^p,k}^o$, of the Lyapunov function (A.27), the first difference along the system dynamics (41) for $k = k_i$ can be expressed as

$$\Delta L_{x^p,k}^o = [(A - B\hat{K}_k)x_k^p + B\hat{K}_k e_k^x]^T[(A - B\hat{K}_k)x_k^p + B\hat{K}_k e_k^x] - x_k^{pT} x_k^p.$$

After simple mathematical manipulation using C-S inequality, Lemma 4.6 and norm, the first difference leads to

$$\begin{aligned}\Delta L^o_{x^p,k} \le & -(1-3\mu)\left\|x_k^p\right\|^2 + 3B_{\max}^2 l_{\tilde{\Theta}_o}\left\|\tilde{\Theta}_k\right\|^2 + 6B_{\max}^2 K_M^2\left\|e_k^x\right\|^2 \\ & +3B_{\max}^2 l_{\tilde{\Theta}_o}^2\left\|\tilde{\Theta}_k\right\|^4 + 3B_{\max}^2\left\|e_k^x\right\|^4.\end{aligned} \tag{A.28}$$

where $\left\|\tilde{K}_k^o x_k^p\right\|^2 \le l_{\tilde{\Theta}_o}\left\|\tilde{\Theta}_k^T Z_k^o Z_k^{oT}\tilde{\Theta}_k\right\| \Big/ \left\|I + Z_k^{o^T} Z_k^o\right\|$ and $\tilde{K}_k^o = K^* - \hat{K}_k^o$ with $\left\|K^*\right\| = K_M$.

Next, the second term, $L^o_{O,k}$, of the Lyapunov function, the first difference is same as in (A.26) of Lemma 5.1.

Now, consider the third term, $L^o_{\tilde{\Theta}1,k} = \tilde{\Theta}_k^T\tilde{\Theta}_k$. The first difference along the dynamics (38) for $k = k_i$, augmented Bellman error (39) and C-S inequality, leads to

$$\begin{aligned}\Delta L^o_{\tilde{\Theta}1,k} \le & -\alpha_V^o(1-3\alpha_V^o)\frac{\tilde{\Theta}_k^T Z_k^o Z_k^{o^T}\tilde{\Theta}_k}{\left\|I + Z_k^{o^T} Z_k^o\right\|} + \alpha_V^o(2+3\alpha_V^o)\frac{\Theta^T(Z_k^o - Z_k^p)(Z_k^o - Z_k^p)^T\Theta}{\left\|I + Z_k^{o^T} Z_k^o\right\|} \\ & +\alpha_V^o(2+3\alpha_V^o)\frac{(F_k^o - F_k^p)(F_k^o - F_k^p)^T}{\left\|I + Z_k^{o^T} Z_k^o\right\|}.\end{aligned}$$

Observing that $0 < \bar{Z}^o_{\min} \le \left\|Z_k^o Z_k^{o^T}\right\| \Big/ \left\|I + Z_k^{o^T} Z_k^o\right\| < 1$, which is ensured by using PE condition, $\left\|Z_k^o - Z_k^p\right\| \le L_Z\left\|e_k^x\right\|$ and $\left\|F_k^o - F_k^p\right\| \le L_F\left\|e_k^x\right\|$ by Lipschitz continuity, the upper bound on the first difference can be represented as

$$\begin{aligned}\Delta L^o_{\tilde{\Theta}1,k} &\le -\alpha_V^o(1-3\alpha_V^o)\bar{Z}^{o^2}_{\min}\left\|\tilde{\Theta}_k\right\|^2 + \alpha_V^o(2+3\alpha_V^o)(L_F^2 + \Theta_M^2 L_Z^2)\left\|e_k^x\right\|^2 \\ &= -\varpi_3\left\|\tilde{\Theta}_k\right\|^2 + \varpi_4\left\|e_k^x\right\|^2,\end{aligned} \tag{A.29}$$

where $\varpi_3 = \alpha_V^o(1-3\alpha_V^o)\bar{Z}^{o^2}_{\min}$ and $\varpi_4 = \alpha_V^o(2+3\alpha_V^o)(L_F^2 + \Theta_M^2 L_Z^2)$ for brevity. Note $\varpi_3 > 0$ and $\varpi_4 > 0$ by selecting the learning gain $0 < \alpha_V^o < 1/3$.

Considering the fifth term, $L^o_{O2,k}$, the first difference using (A.26) can be expressed as

$$\begin{aligned}\Delta L^o_{O2,k} &= (L^o_{O,k+1})^2 - (L^o_{O,k})^2 = \Delta L^{o^2}_{O,k} + 2\Delta L^o_{O,k} L^o_{O,k} \\ &\le -2\Lambda_3^2 \varpi_1 (1-\varpi_1) \left\| e^x_k \right\|^4 - 2\varpi_2 (1-\varpi_2) \left\| \tilde{\psi}_k \right\|^4,\end{aligned} \tag{A.30}$$

where $(1-\varpi_1) > 0$ and $(1-\varpi_2) > 0$ by definition.

Finally, the first difference of the last term, $L^o_{\tilde{\Theta}_2,k} = (\tilde{\Theta}^T_k \tilde{\Theta}_k)^2$ using the first difference (A.29), becomes

$$\Delta L^o_{\tilde{\Theta}_2,k} \le -(\varpi_3(2-\varpi_3)/2) \left\| \tilde{\Theta}_k \right\|^4 + (\varpi_4{}^2(2-\varpi_3)/2\varpi_3) \left\| e^x_k \right\|^4, \tag{A.31}$$

where $2-\varpi_3 > 0$ since $0 < \alpha^o_V < 1/3$.

At the final step, the overall first difference $\Delta L^o_k = \Delta L^o_{x^p,k} + \Sigma_{o1} \Delta L^o_{O,k} + \Sigma_{\Theta 1} \Delta L^o_{\tilde{\Theta}1,k}$ $+\Sigma_{o2} \Delta L^o_{O2,k} + \Sigma_{\tilde{\Theta}2} \Delta L^o_{\tilde{\Theta}_2,k}$, from (A.28), (A.26), (A.29), (A.30), (A.31), and recalling the definition of $\Sigma_{o1}, \Sigma_{\Theta 1}, \Sigma_{\Theta 2}$, and Σ_{o2}, becomes

$$\begin{aligned}\Delta L^o_k \le & -(1-3\mu) \left\| x^p_k \right\|^2 - \left(6B^2_{\max} K^2_M + \Sigma_{\Theta 1} \varpi_4\right) \left\| e^x_k \right\|^2 - \Sigma_{o1} \varpi_2 \left\| \tilde{\psi}_k \right\|^2 \\ & - B^2_{\max} l_{\tilde{\Theta}_o} \left\| \tilde{\Theta}_k \right\|^2 - \left(6B^2_{\max} \varpi_3 + \Sigma_{\tilde{\Theta}2} \varpi_4{}^2 (2-\varpi_3)\right) / 2\varpi_3 \left\| e^x_k \right\|^4 \\ & - 2\Sigma_{o2} \varpi_2 (1-\varpi_2) \left\| \tilde{\psi}_k \right\|^4 - B^2_{\max} l^2_{\tilde{\Theta}_o} \left\| \tilde{\Theta}_k \right\|^4\end{aligned} . \tag{A.32}$$

From (A.32), the first difference of the Lyapunov function $\Delta L^o_k < 0$. Therefore, by Lyapunov theorem (Jagannathan, 2006), the Lyapunov function is a non-increasing function i.e., $L^o_{k_i+1} < L^o_{k_i}$..

Case 2. *During the inter-event times ($k_i < k < k_{i+1}$)*

Consider the Lyapunov function L^o_k same as in Case 1. The first difference of the first term, along the system dynamics (41) can be written as

$$\Delta L^{o}_{x^p,k} = ((A - B\hat{K}_k)x_k^p + B\hat{K}_k e_k^x + B\hat{K}_k e_k^{o,ET})^T((A - B\hat{K}_k)x_k^p + B\hat{K}_k e_k^x + B\hat{K}_k e_k^{o,ET}) - x_k^{p^T} x_k^p$$

$$\leq -(1-4\mu)\left\|x_k^p\right\|^2 + 4\left\|B\tilde{K}_k x_k^p\right\|^2 + 4\|B\|^2\left\|\hat{K}_k\right\|^2\left\|e_k^x\right\|^2 + 4\|B\|^2\left\|\hat{K}_k\right\|^2\left\|e_k^{o,ET}\right\|^2. \quad \text{(A.33)}$$

Recalling the triggering condition (42) one can reach at

$$\begin{aligned}\Delta L^{o}_{x^p,k} \leq & -(1-2\Gamma^{o}_{ET})(1-4\mu)\left\|x_k^p\right\|^2 + 4B^2_{\max} l_{\tilde{\Theta}_o}\left\|\tilde{\Theta}_k\right\|^2 + 4B^2_{\max} l^2_{\tilde{\Theta}_o}\left\|\tilde{\Theta}_k\right\|^4 \\ & +2(4B^2_{\max}K_M^2 + \Gamma^{o}_{ET}(1-4\mu))\left\|e_k^x\right\|^2 + 4B^2_{\max}\left\|e_k^x\right\|^4.\end{aligned} \quad \text{(A.34)}$$

Moreover, the first difference $\Delta L^{o}_{O,k}$ remains same as (A.26) in Lemma 5.1. Moving on, the first differences $\Delta L^{o}_{\tilde{\Theta}1,k}$ and $\Delta L^{o}_{\tilde{\Theta}2,k}$ along the observer QFE parameter estimation error dynamics (38) for $k_i < k < k_{i+1}$ are given by

$$\Delta L^{o}_{\tilde{\Theta}1,k} = \tilde{\Theta}^T_{k+1}\tilde{\Theta}_{k+1} - \tilde{\Theta}^T_k\tilde{\Theta}_k = 0 \text{ and } \Delta L^{o}_{\tilde{\Theta}2,k} = 0. \quad \text{(A.35)}$$

Finally, the first difference of $L^{o}_{O2,k}$ is same as in (A.30). At the final step, the overall first difference $\Delta L^{o}_k = \Delta L^{o}_{x^p,k} + \Sigma_{o1}\Delta L^{o}_{O,k} + \Sigma_{\Theta 1}\Delta L^{o}_{\tilde{\Theta}1,k} + \Sigma_{o2}\Delta L^{o}_{O2,k} + \Sigma_{\tilde{\Theta}2}\Delta L^{o}_{\tilde{\Theta}_2,k}$, using (A.26), (A.30), (A.34), (A.35), and definitions of Σ_{o1} and Σ_{o2}, becomes

$$\begin{aligned}\Delta L^{o}_k \leq & -\left(1-2\Gamma^{o}_{ET}\right)\left(1-4\mu\right)\left\|x_k^p\right\|^2 - B^2_{\max}\left\|e_k^x\right\|^4 - \Sigma_{o1}\varpi_2\left\|\tilde{\psi}_k\right\|^2 \\ & -2\Sigma_{o2}\varpi_2\left(1-\varpi_2\right)\left\|\tilde{\psi}_k\right\|^4 - \left(4B^2_{\max}K_M^2 + \Gamma^{o}_{ET}\left(1-4\mu\right)\right)\left\|e_k^x\right\|^2 + B^{o}_{\tilde{\Theta},k_i+1},\end{aligned} \quad \text{(A.36)}$$

where $B^{o}_{\tilde{\Theta},k_i+1} = 4B^2_{\max} l_{\tilde{\Theta}_o}\left\|\tilde{\Theta}_{k_i+1}\right\|^2 + 2(2B^2_{\max} + \Gamma^{o}_{ET}(1-4\mu))l^2_{\tilde{\Theta}_o}\left\|\tilde{\Theta}_{k_i+1}\right\|^4$, $0 < \Gamma^{o}_{ET} < 1/2$, and $0 < \mu < 1/4$. Since, $\tilde{\Theta}_{k_i+1} = \tilde{\Theta}_k = \tilde{\Theta}_{k_{i+1}}$ for $k_i < k < k_{i+1}$ from (A.35), the expression for $B^{o}_{\tilde{\Theta},k_i+1}$ in (A.36) uses $\tilde{\Theta}_{k_i+1}$. Therefore, $B^{o}_{\tilde{\Theta},k_i+1}$ is a piecewise constant function. Therefore, $B^{o}_{\tilde{\Theta},k_i+1}$ is a piecewise constant function.

From (A.36), the Lyapunov first difference $\Delta L_k^0 < 0$ as long as,

$$\left\| x_k^p \right\| > \sqrt{B_{\tilde{\Theta},k_i+1}^o \Big/ (1-2\Gamma_{ET}^o)(1-4\mu)} \equiv B_{x^p 2,k_i+1}^o \text{ or}$$

$$\left\| e_k^x \right\| > \max\left\{ \sqrt{B_{\tilde{\Theta},k_i+1}^o \Big/ \left(4B_{\max}^2 K_M^2 + \Gamma_{ET}^o (1-4\mu)\right)}, \sqrt[4]{B_{\tilde{\Theta},k_i+1}^o \Big/ B_{\max}^2} \right\} \equiv B_{e^x 2,k_i+1}^o \text{ or}$$

$$\left\| \tilde{\psi}_k \right\| > \max\left\{ \sqrt{B_{\tilde{\Theta},k_i+1}^o \Big/ \Sigma_{o1}\varpi_2}, \sqrt[4]{B_{\tilde{\Theta},k_i+1}^o \Big/ 2\Sigma_{o2}\varpi_2 (1-\varpi_2)} \right\} \equiv B_{\tilde{\psi} 2,k_i+1}^o.$$

A common bound for x_k^p, e_k^x, and $\tilde{\psi}_k$ can be selected as $B_{o,k_i+1}^M = \max\{B_{x^p 2,k_i+1}^o, B_{e^x 2,k_i+1}^o, B_{\tilde{\psi} 2,k_i+1}^o\}$. From (A.36), by Lyapunov theorem, x_k^p, e_k^x, $\tilde{\psi}_k$, and $\tilde{\Theta}_k$ are bounded during the inter-event time. Further, x_k^p, e_k^x, and $\tilde{\psi}_k$ will converge to the bound B_{o,k_i+1}^M in finite time and $\tilde{\Theta}_k$ remains constant. It follows that, the Lyapunov function $L_k^o \to B_{L^o,k}^o$, $k_i < k < k_{i+1}$ and stay within the bound where

$$B_{L^o,k}^o = (B_{o,k_i+1}^M)^2 + \Sigma_{o1}(B_{o,k_i+1}^M)^2 + \Sigma_{\Theta 1}(\tilde{\Theta}_{k_i+1})^2 + \Sigma_{o2}(B_{o,k_i+1}^M)^4 + \Sigma_{\tilde{\Theta} 2}(\tilde{\Theta}_{k_i+1})^4.$$

The bound $B_{L^o,k}^o$ is computed from (A.27) using the upper bound for x_k^p, e_k^x, $\tilde{\psi}_k$ and $\tilde{\Theta}_k$.

Now, defining the upper bound function

$$h_o(k) = \max(L_k^o, B_{L^o,k}^o),$$

it is clear that $h_o(k) \geq L_k^o$ for all $k \in \mathbb{N}$. Further, $\Delta L_k^o < 0$, $k = k_i$ implies $L_{k_i+1}^0 < L_{k_i}^0$ and, hence, the QFE parameter estimation error $\tilde{\Theta}_{k_i+1} < \tilde{\Theta}_{k_i}$ and, from (A.35) in Case II, $\tilde{\Theta}_{k_i+1} = \tilde{\Theta}_k = \tilde{\Theta}_{k_{i+1}}$, $k_i < k < k_{i+1}$. This implies $\tilde{\Theta}_{k_{i+1}} < \tilde{\Theta}_{k_i}$. Therefore, $\tilde{\Theta}_{k_i} \to 0$ as event sampled instants $k_i \to \infty$ since $\tilde{\Theta}_{k_i} \to 0$. It follows that, $B_{L^o,k}^o \to 0$ as $k_i \to \infty$.

Therefore, $L_k^o \to 0$ as $k_i \to \infty$ or alternatively, $k \to \infty$. This implies $h_o(k) \to 0$ as $k \to \infty$.Finally, using the above results, it is routine to check, $\hat{Q}_k \to Q_k^*$ and $u_k \to u_k^*$ as $k \to \infty$. ■

II. ADAPTIVE NEURAL NETWORK BASED EVENT-TRIGGERED CONTROL OF SINGLE-INPUT SINGLE-OUTPUT NONLINEAR DISCRETE TIME SYSTEMS

Avimanyu Sahoo, Hao Xu and S. Jagannathan

Abstract *— This paper presents a novel adaptive neural network (NN) control of single-input and single-output (SISO) uncertain nonlinear discrete time systems under event sampled NN inputs. In this control scheme, the feedback signals are transmitted and the NN weights are tuned in an aperiodic manner at the event sampled instants. After reviewing the NN approximation property with event sampled inputs, an adaptive state estimator (SE), consisting of linearly parameterized NNs, is utilized to approximate the unknown system dynamics in an event sampled context. The SE is viewed as a model and its approximated dynamics and the state vector, during any two events, are utilized for the event-triggered controller design. An adaptive event-triggering condition is derived by using both the estimated NN weights and a dead-zone operator to determine the event sampling instants. This condition both facilitates the NN approximation and reduces the transmission of feedback signals. The ultimate boundedness (UB) of both the NN weight estimation error and system state vector is demonstrated via Lyapunov approach. As expected, during an initial online learning phase, events are observed more frequently. Over time with the convergence of the NN weights, the inter-event times increase thereby lowering the number of triggered events. These claims are illustrated via simulation results.*

Index Terms - Adaptive control, event-triggered control, function approximation, neural network control.

1. INTRODUCTION

Traditional periodic transmission of feedback control signals in a closed-loop networked environment requires a higher network bandwidth. Event-triggered control (ETC) [1]-[16], on the other hand, is emerged recently as an alternate method to reduce the network communication and controller execution. In ETC, the aperiodic sampling of system state vector is proven to be advantageous computationally over periodic sampled control schemes [1].

The ETC technique allows the system errors to increase to a predefined threshold before transmitting the feedback signals. The threshold is designed to both avoid instability and meet a certain desired performance. Therefore, the transmissions of the feedback signals and control input are reduced while achieving a desired control performance. These transmission instants are usually referred to as *event sampled instants or* simply *event-triggered instants* [2]. The condition under which a decision is made to transmit the feedback and control signals is known as *event-triggering condition* [2]. The event-triggering condition is normally a function of the system state error which is referred to as event-trigger error [2]-[13] along with a state dependent threshold.

In an earlier work [2] on ETC, the authors assumed input-to-state stability (ISS) of the system with respect to the event-trigger error for designing an event-triggering condition. It was shown that the event-based controller ensured the asymptotic stability of the system with reduced computation. Later, various other ETC schemes [1], [4]-[16] are developed for both linear and nonlinear systems. A majority of these ETC schemes are implemented by using a zero-order-hold (ZOH) [2]-[5] in order to maintain both the last transmitted state vector and control input until the next transmission.

An alternate to the ZOH scheme is the model-based scheme [7]-[9], [11], [14]-[15] where the state vector from a model is used to generate the control input within any two event-triggered instants. The model-based approach is shown to reduce network traffic more than a ZOH-based scheme at the expense of additional computation due to the model. However, in all the ETC effort [2]-[11], the system dynamics are considered available a priori while a small bounded uncertainty can be tolerated [8]. In contrast, in our preliminary work [14]-[15], adaptive model-based schemes both for uncertain linear systems and partially unknown nonlinear systems, respectively, were introduced.

From the stability point of view and to account for the aperiodic transmissions of the feedback signals, several closed-loop modelling techniques are also presented. A representative list includes the piece-wise linear system model [11], the perturbed system model [8], the hybrid and impulsive [11] dynamical system models. All these modelling approaches utilized the Lyapunov method or its extension for the stability analysis and to design the event-triggering condition.

In this paper, an adaptive model-based ETC scheme for a nonlinear discrete-time system in Brunovsky canonical form is presented. Both the internal dynamics and the control coefficient function are considered unknown. By using the approximation property of neural networks (NN) [20], in an event sampled context, an adaptive state estimator (SE) is designed. The adaptive SE serves as a model of the system and both approximates the system dynamics and estimates the state vector. The approximated system dynamics and the estimated state vector are subsequently utilized for generating the control input, during any two event sampled instants.

A novel event-triggering condition is derived using the Lyapunov method of stability. The threshold in the event-triggering condition is designed as a function of both the NN weight estimates and the system state vector. Thus, the threshold becomes adaptive unlike the traditional threshold conditions [2]-[7] which are functions of system state vector alone. This modified adaptive event-triggering condition not only ensures the function approximation by using a non-periodic weight update law but also the stability. The event-triggering condition further uses a dead-zone operator to prevent the unnecessary triggering of events, due to the NN reconstruction error, once the system state is inside the ultimate bound.

The contributions of this paper include: a) the event sampled NN approximation with model state vector, b) the development of a novel model-based adaptive NN ETC scheme, c) an aperiodic tuned NN-based state estimator (SE) or model, and, d) an adaptive event-triggering condition to ensure the stability and convergence of NN weight estimates.

The completely uncertain system dynamics make the event-triggering condition design different from the traditional one [2]-[8] including partially unknown dynamics in [15] The stability of the event-triggered closed-loop system is proven by using the idea of switched systems as discussed in [10], [17]. The Lyapunov function is allowed to increase during the inter-event times but bounded. It is shown that the bound for the Lyapunov function during inter-event times converge to the ultimate value with events occurring. This enables the proposed NN-based adaptive event-triggered scheme to ensure stability in the presence of significant level of dynamic uncertainty. It also reduces the network traffic with fewer numbers of triggered events when compared to a traditional discrete-time system.

The remaining part of the paper is organized as follows. Section 2 revisits the event-based approximation and formulates the problem for the ETC of uncertain dynamical systems. Section 3 details the design procedure for the NN-based adaptive ETC. The stability is claimed in Section 4. Before concluding in Section 5, the simulation results are presented in Section 6. The Appendix details the proofs for the lemmas and theorems.

2. BACKGROUND

This section presents a brief background on the traditional ETC and formulates the problem for adaptive ETC.

2.1 BACKGROUND ON ETC

Consider a controllable nonlinear uncertain discrete-time system in Brunovsky canonical form given by

$$\begin{aligned} x_{1,k+1} &= x_{2,k}, \\ x_{2,k+1} &= x_{3,k}, \\ &\vdots \\ x_{n,k+1} &= f(x_k) + g(x_k)u_k, \\ y_k &= x_{1,k}, \end{aligned} \tag{1}$$

where $x_k = [x_{1,k} \quad x_{2,k} \quad \cdots \quad x_{n,k}]^T \in \Re^n$, $u_k \in \Re$, and $y_k \in \Re$ denote the state vector, the input and the output of the system. The internal dynamics and the control coefficient function, $f : \Re^n \to \Re$ and $g : \Re^n \to \Re$, respectively, are unknown nonlinear smooth functions. The system is considered to be feedback linearizable [21] in the sense that there exists a diffeomorphism to transform the system into a linear form**.**

The system (1) can be written in simplified form as

$$x_{k+1} = Ax_k + Bf(x_k) + Bg(x_k)u_k, \tag{2}$$

where $A = \begin{bmatrix} 0 & 1 & \cdots & 0 \\ 0 & 0 & \ddots & 0 \\ \vdots & \vdots & \ddots & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \in \Re^{n\times n}$ and $B = \begin{bmatrix} 0 & 0 & \cdots & 1 \end{bmatrix}^T \in \Re^n$.

The system dynamics (2) can be rewritten in a compact form as

$$x_{k+1} = Ax_k + B\bar{F}(x_k)\bar{u}_k, \tag{3}$$

where $\bar{F}(x_k) = [\, f(x_k) \quad g(x_k)] \in \Re^{1\times 2}$ and $\bar{u}_k = [1 \quad u_k]^T \in \Re^2$ are the augmented system dynamics and input vector, respectively. These augmented forms are utilized in the model development and controller design. To design a controller by using feedback linearization, the following assumption is required.

Assumption 1[18]: The nonlinear function $g(x_k)$ is lower bounded, i.e., $0 < g_{\min} \le \left| g\left(x_k\right) \right|$ where $g_{\min}$ is a known positive constant and $|\bullet|$ denotes the absolute value.

For the system (1), under complete knowledge of system dynamics, a feedback linearizable controller of the following form

$$u_{d_k} = (-f(x_k) + v_k)/g(x_k), \tag{4}$$

yields an asymptotically stable closed-loop system. The closed-loop dynamics can be written as

$$x_{k+1} = A_c x_k, \tag{5}$$

where u_{d_k} is the ideal control input. The stabilizing control input is given by $v_k = Kx_k$ where $K = \begin{bmatrix} K_1 & K_2 & \cdots & K_n \end{bmatrix}$ is the control gain vector. The control gain vector $K \in \Re^n$ can be designed to ensure A_c is Schur via suitable pole placement design. The closed-loop system matrix can be written as $A_c = A + BK = \begin{bmatrix} 0 & 1 & \cdots & 0 \\ 0 & 0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ K_1 & K_2 & \cdots & K_n \end{bmatrix} \in \Re^{n\times n}$. For the class of systems given by (1), any nonlinear controller can also be utilized which renders asymptotic stability of the system. The ideal controller (4) needs time-based periodic

sampled system state x_k for implementation along with $f(x_k)$ and $g(x_k)$. In contrast, our main objective in this paper is to implement the controller (4) in the event sampled context without the knowledge of the system dynamics.

In the case of traditional ETC, the system state vector x_k is transmitted to the controller only at the event sampled instants. Define a subsequence $\{k_i\}_{i=1}^{\infty}$ of the discrete sequence of time instants $k \in \mathbb{N}$, referred to as event sampled instants. The events are triggered at k_i, $\forall\, i \in \mathbb{N}$ with the first event occurring at the time instant $k_0 = 0$. The system state vector, x_{k_i}, is transmitted through the communication network and held by a ZOH till the next transmission at k_{i+1}. The last held state, x_{k_i} for $k_i \le k < k_{i+1}$ at the ZOH is piecewise constant and used for the controller implementation.

The event sampled instants are determined at the trigger mechanism by evaluating the event-trigger error against the threshold value. The deviation between x_k and the last transmitted state x_{k_i} is usually referred to as the event-trigger error, e_k^{ET}. This is represented as

$$e_k^{ET} = x_k - x_{k_i}, \quad k_i \le k < k_{i+1},\ i \in \mathbb{N}. \tag{6}$$

Though event-triggering condition is evaluated periodically at all $k \in \mathbb{N}$, the state vector is transmitted to the controller only at the event sampled instants determined by the violation of the event-triggering condition.

Now, to implement the controller (4), the unknown system dynamics, $f(x_k)$ and $g(x_k)$, must be approximated by using event sampled system state vector, x_{k_i}. Therefore,

the universal approximation property of the NN is revisited for event-based sampling in the next subsection.

2.2 PROBLEM FORMULATION

The problem of ETC is formulated in this subsection by addressing event sampled NN approximation and transmission of state vector.

2.2.1 Event Sampled Neural Network Approximation. According to the universal approximation property [20] of the NN, a nonlinear smooth function $h(x_k) \in \Re^n$ can be approximated in compact set for all $x_k \in \Omega_x \subset \Re^n$. A linearly parameterized NN [20] with two hidden layers can be used for the purpose. The two layer NN can consists of a layer of randomly assigned constant weights, V_h, in the input layer and tunable weight matrix, W_h, in the output layer. It has been proven that by randomly selecting the input layer weights, the activation function forms a stochastic basis [20]. Thus, the NN approximation property holds [20] for all inputs x_k belong to a compact set $\Omega_x \subset \Re^n$. The function $h(x_k) \in \Re^n$ with the linearly parametrized NN can be represented as

$$h(x_k) = W_h^T \psi_h(V_h^T x_k) + \varepsilon_h(x_k), \tag{7}$$

where $W_h \in \Re^{l \times n}$ is the NN target weight matrix. The randomly assigned input weight matrix is denoted by $V_h \in \Re^{a \times l}$ and $\psi_h(\bullet) \in \Re^l$ is the activation function vector. The NN reconstruction error, the number of hidden layer neurons, and the number of inputs are denoted by $\varepsilon_h(x_k) \in \Re^n$, l, and a, respectively. So far in the literature, the universal NN approximation property considers the availability of x_k at all-time instants k .

In the case of an ETC, the approximation of the function at event sampled instants $h(x_{k_i})$ can be expressed as [16]

$$h(x_{k_i}) = W_h^T \psi_h(V_h^T x_{k_i}) + \varepsilon_h(x_{k_i}), \tag{8}$$

where $\psi_h(V_h^T x_{k_i}) \in \Re^l$ is the activation function with event sampled state vector, x_{k_i}. The reconstruction error at event sampled instants is given by $\varepsilon_h(x_{k_i}) \in \Re^n$. Note that, the approximations (7) and (8) become equal if the events are triggered at all-time instants. Since the events are occurring in an aperiodic manner, the function $h(x_k)$ for $k_i \le k < k_{i+1}$ can be expressed as

$$h(x_k) = W_h^T \psi_h(V_h^T x_{k_i}) + \varepsilon_{h,e}(x_{k_i}, e_k^{ET}),\ k_i \le k < k_{i+1},\ i \in \mathbb{N}, \tag{9}$$

where $\varepsilon_{h,e}(x_{k_i}, e_k^{ET})$ is the event sampled reconstruction error computed next.

Consider the periodic approximation of the function $h(x_k)$ as in (7). By adding and subtracting $\psi_h(V_h^T x_{k_i})$ and definition (6), it can be rewritten as

$$\begin{aligned} h(x_k) &= W_h^T \psi_h(V_h^T x_k) + W_h^T \psi_h(V_h^T x_{k_i}) - W_h^T \psi_h(V_h^T x_{k_i}) + \varepsilon_h(x_k) \\ &= W_h^T \psi_h(V_h^T x_{k_i}) + W_h^T \left(\psi_h(V_h^T (x_{k_i} + e_k^{ET})) - \psi_h(V_h^T x_{k_i})\right) \\ &\quad + \varepsilon_h(x_{k_i} + e_k^{ET}),\ \ k_i \le k < k_{i+1},\ \ \forall i \in \mathbb{N}. \end{aligned} \tag{10}$$

Comparing (9) and (10) the event sampled NN reconstruction error is given by $\varepsilon_{h,e}(x_{k_i}, e_k^{ET}) = \varepsilon_h(x_{k_i} + e_k^{ET}) + W_h^T \left(\psi_h(V_h^T (x_{k_i} + e_k^{ET})) - \psi_h(V_h^T x_{k_i})\right)$. The event sampled NN reconstruction error, $\varepsilon_{h,e}(x_{k_i}, e_k^{ET})$, is a function of the traditional NN reconstruction error $\varepsilon_h(x_k) = \varepsilon_h(x_{k_i} + e_k^{ET})$ as in (7) and an additional error due to event sampled input, i.e., $W_h^T \left(\psi_h(V_h^T (x_{k_i} + e_k^{ET})) - \psi_h(V_h^T x_{k_i})\right)$. This additional error is a function of event

sampled state vector, x_{k_i}, and the event-trigger error e_k^{ET}. Therefore, to approximate a function with a desired level of accuracy in an ETC context, the event-trigger error, e_k^{ET}, must be kept small. This can be achieved by designing a suitable event-triggering condition. Higher is the number of event sampled instants, better will be the NN approximation. However, this will increase the number of transmissions leading to higher network bandwidth usage.

The NN estimation of the function $\hat{h}(x_k)$ for $k_i \le k < k_{i+1}, \forall i \in \mathbb{N}$ can be written as

$$\begin{aligned} \hat{h}(x_k) &= \hat{W}_{h,k}^T \psi_h(V_h^T x_k) \\ &= \hat{W}_{h,k}^T \psi_h(V_h^T x_{k_i}) + \hat{W}_{h,k}^T \left(\psi_h(V_h^T (x_{k_i} + e_k^{ET})) - \psi_h(V_h^T x_{k_i}) \right), \end{aligned} \tag{11}$$

where $\hat{W}_{h,k} \in \Re^{l \times n}$ is the NN weight estimate. The second term $\hat{W}_{h,k}^T(\psi_h(V_h^T(x_{k_i} + e_k^{ET})) - \psi_h(V_h^T x_{k_i}))$ is an additional error in estimation and a function of event-trigger error.

It is important to mention here that the event-based aperiodic transmission precludes the traditional periodic NN weight update [20]. The NN weights must be tuned in an aperiodic manner only at the event sampled instants, $k = k_i$ with the latest measuring state vector. This, further, requires a suitable event-triggering condition.

From the above discussion, the accuracy of NN approximation, the reduction in transmissions, and the system stability depend upon the event-triggering condition. Thus, a trade-off must be reached through a careful design of the event-triggering condition. As a solution, the threshold of the event-triggering condition is made adaptive in contrast with the fixed threshold utilized in traditional ETC design with known dynamics [2]-[5].

An alternate to this ZOH based technique is the model-based approach and discussed next.

2.2.2 Model-base ETC. The structure of a model-based event-triggered control (MBETC) scheme [7]-[9], [11] is illustrated in Figure 1.

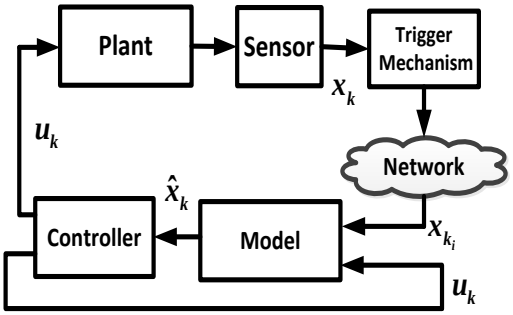

Figure 1. Structure of the traditional MBETC system**.**

Traditionally, a system model (known a priori) generates the state vector between the event sampled instants. The model state vector is subsequently used by the controller to update the control input periodically in contrast with a ZOH ETC. The event sampled instants are determined by the deviation of the model state from the measured system state vector due to model uncertainty or disturbance. The measured system state vector is transmitted at the event sampled instants to reinitialize the model state vector.

The event-trigger error for a MBETC scheme can be redefined as the difference between the measured system state and the model state vector. It is given by

$$e_k^s = x_k - \hat{x}_k,\ k_i \le k < k_{i+1}\ ,\ \forall i \in \mathbb{N}\,, \tag{12}$$

where $\hat{x}_k \in \Re^n$ is the model state vector. The reinitialized model state vector at the trigger instants can be represented as

$$\hat{x}_k = x_k, \quad k = k_i, \ \forall i \in \mathbb{N}, \tag{13}$$

and then it evolves with model dynamics during the inter-event times for $k_i < k < k_{i+1}$. Since the system dynamics in (1) are uncertain, the traditional model-based ETC framework cannot be directly used. This requires an adaptive NN scheme to construct the model or SE. Further, the model dynamics must also be approximated in the MBETC context similar to the ZOH based case as discussed before. The detailed design procedure is presented next.

3. MODEL BASED ADAPTIVE ETC DESIGN

The adaptive MBETC scheme for an uncertain nonlinear discrete-time system is proposed in this section. We assume a communication network between the sensor and controller but without packet losses and delays. This assumption is consistent with the ETC literature [9], [11] for the purpose of controller design.

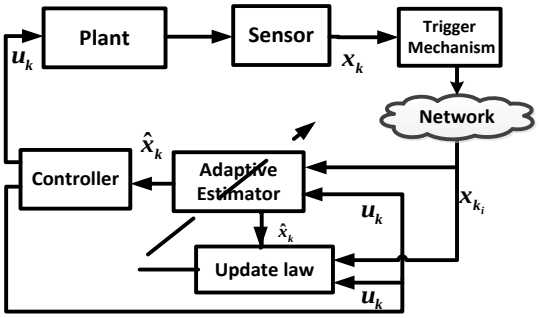

Figure 2. Structure of the adaptive METC system.

The structure of the traditional MBETC, shown in Figure 1, is modified for an adaptive MBETC and shown in Figure 2. A NN-based adaptive model or SE is included not only to estimate the state vector but also to approximate the unknown system dynamics. An adaptive event-triggering condition is also proposed using the SE's estimated NN weights and the system state. Therefore, a mirror SE at the trigger mechanism is used to evaluate the event-triggering condition. This mirror SE estimates the NN weights locally at the trigger mechanism to avoid the transmission of the NN weight estimates through the communication network. The mirror SE operates in synchronism with the SE at the controller. At the violation of the triggering condition at $k = k_i$, the system state x_{k_i} and x_{k_i-1}, are transmitted together. The received state vectors

are used to update the NN weights at trigger instants in an aperiodic manner. Then, the event-trigger error, e_k^s, in (12) is reset to zero for the next cycle of triggering. The detailed design procedure for the NN-based adaptive MBETC scheme is presented next.

3.1 ADAPTIVE ESTIMATOR AND CONTROLLER DESIGN

The dynamics of the adaptive SE can be expressed as

$$\hat{x}_{k+1} = A\hat{x}_k + B\hat{f}(\hat{x}_k) + B\hat{g}(\hat{x}_k)u_k, \;\; k_i \le k < k_{i+1}, \; \forall i \in \mathbb{N}, \tag{14}$$

where $\hat{x}_k = [\hat{x}_{1,k} \quad \hat{x}_{2,k} \quad \cdots \quad \hat{x}_{n,k}]^T \in \Re^n$ represents the estimated state vector. The functions $\hat{f}(\hat{x}_k) \in \Re$ and $\hat{g}(\hat{x}_k) \in \Re$ represent the approximation of the nonlinear functions $f(x_k) \in \Re$ and $g(x_k) \in \Re$, respectively. The system state vector, x_k, is available intermittently only at $k = k_i$. Thus, the approximation of the nonlinear functions are express as $\hat{f}(\hat{x}_k)$ and $\hat{g}(\hat{x}_k)$ with the estimated state vector, $\hat{x}_k$. Further, as proposed, the SE sate vector is reinitialized as in (13).

The dynamics of the SE in (14), in an augmented form as in (3), for both inter-event times and trigger instants using (13), can be represented as

$$\hat{x}_{k+1} = \begin{cases} A\hat{x}_k + B\hat{\bar{F}}(\hat{x}_k)\bar{u}_k, & k_i < k < k_{i+1}, \\ Ax_k + B\hat{\bar{F}}(x_k)\bar{u}_k, & k = k_i, \end{cases} \tag{15}$$

where $\hat{\bar{F}}(\hat{x}_k) = [\hat{f}(\hat{x}_k) \quad \hat{g}(\hat{x}_k)]$ and $\hat{\bar{F}}(x_k) = [\hat{f}(x_k) \quad \hat{g}(x_k)]$.

Consider the augmented system dynamics (3). The nonlinear function, $\bar{F}(x_k)$, can be approximated in the event-triggered context, similar to (8) and (9), using the SE state $\hat{x}_k$ as input to the NN. Hence, $\bar{F}(x_k)$ can be expressed as

$$\bar{F}\left(x_k\right) = W^T \Phi(\hat{\bar{x}}_k) + \Xi_e(\hat{x}_k, e_k^s), \quad k_i \le k < k_{i+1}, \ \forall i \in \mathbb{N}, \tag{16}$$

where $W = [W_f^T \quad W_g^T]^T \in \Re^{2l \times 1}$ is the unknown target NN weight matrix with $W_f \in \Re^l$ and $W_g \in \Re^l$ represents the target weights for $f(x_k)$ and $g(x_k)$. The event-based activation function matrix is denoted by $\Phi\left(\hat{\bar{x}}_k\right) = \begin{bmatrix} \varphi_f(V_f^T \hat{x}_k) & 0_{l\times 1} \\ 0_{l\times 1} & \varphi_g(V_g^T \hat{x}_k) \end{bmatrix} \in \Re^{2l \times 2}$ where $V_f \in \Re^{a \times l}$ and $V_g \in \Re^{a \times l}$ are randomly assigned constant weights at the input layers, and $\varphi_f(\bullet)$ and $\varphi_g(\bullet)$ represent the NN activation functions for f and g, respectively. The input matrices can be selected as $V_f = V_g = V$. Then, $\hat{\bar{x}}_k = V_f^T \hat{x}_k = V_g^T \hat{x}_k = V^T \hat{x}_k$.

The event-based reconstruction error using SE state vector is denoted by $\Xi_e(\hat{x}_k, e_k^s) = \Xi(\hat{x}_k + e_k^s) + W^T(\Phi(V^T(\hat{x}_k + e_k^s)) - \Phi(V^T \hat{x}_k))$ where $\Xi(\hat{x}_k + e_k^s) = \Xi(x_k)$ $= [\varepsilon_f(x_k) \quad \varepsilon_g(x_k)] \in \Re^{1\times 2}$ is the traditional reconstruction error in augmented form. The reconstruction errors $\varepsilon_f(x_k)$ and $\varepsilon_g(x_k)$ are the errors for the function f and g, respectively. The additional error term, as in the case of ZOH based approximation in (9), is given by $W^T \Phi((V^T(\hat{x}_k + e_k^s)) - \Phi(V^T \hat{x}_k))$.

The actual NN estimation of the function, $\hat{\bar{F}}(x_k)$, with SE state, $\hat{x}_k$, can be written similar to (11) for $k_i \le k < k_{i+1}$ as

$$\hat{\bar{F}}\left(x_k\right) = \hat{W}_k^T \Phi(\bar{x}_k) = \hat{W}_k^T \Phi(\hat{\bar{x}}_k) + \hat{W}_k^T \left(\Phi(V^T(\hat{x}_k + e_k^s)) - \Phi(V^T \hat{x}_k)\right), \tag{17}$$

where $\hat{W}_k = [\hat{W}_{f,k}^T \quad \hat{W}_{g,k}^T]^T \in \Re^{2l \times 1}$ represents the estimated NN weight vector and $\Phi\left(\bar{x}_k\right) = \begin{bmatrix} \varphi_f(V^T x_k) & 0_{l\times 1} \\ 0_{l\times 1} & \varphi_g(V^T x_k) \end{bmatrix}$ is the augmented activation function with $\bar{x}_k = V^T x_k$.

Remark 1: The error term $\Phi(V^T(\hat{x}_k + e_k^s)) - \Phi(V^T\hat{x}_k)$, both in (16) and (17), is the result of the model state $\hat{x}_k$ as input to the activation function during $k_i < k < k_{i+1}$ instead of system state x_k . In the case of a traditional NN based model [20], where the system state is used periodically, this error is not present. Since the activation functions are smooth functions, this error can be represented in terms of event-trigger error, e_k^s by using the Lipschitz continuity as given next.

Assumption 2 [20]: The target weight vector W, the NN activation function $\Phi(\bullet)$ and the reconstruction error $\Xi(\bullet)$ are bounded above satisfying $\|W\| \le W_{\max}$, $\|\Phi(\bullet)\| \le \Phi_{\max}$ and $\|\Xi(\bullet)\| \le \Xi_{\max}$ where $W_{\max}$, $\Phi_{\max}$, and $\Xi_{\max}$ are positive constants.

Assumption 3: The NN activation function $\Phi(\bar{x}_k)$ is Lipschitz continuous on a compact set for all $x_k \in \Omega_x$. Then, there exists a constant $L > 0$ such that $\| \Phi(\bar{x}_k) - \Phi(\hat{\bar{x}}_k) \| \le L \| \bar{x}_k - \hat{\bar{x}}_k \| \le L_\Phi \| e_k^s \|$ are satisfied where $L_\Phi = L\|V\|$ is a constant.

The SE dynamics (15) by NN approximation can be expressed as

$$\hat{x}_{k+1} = \begin{cases} A\hat{x}_k + B\hat{W}_k^T\Phi(\hat{\bar{x}}_k)\bar{u}_k, & k_i < k < k_{i+1}, \\ Ax_k + B\hat{W}_k^T\Phi(\bar{x}_k)\bar{u}_k, & k = k_i. \end{cases} \tag{18}$$

The event-based control input with the estimated SE state vector, $\hat{x}_k$, and the SE dynamics (15), can be represented as

$$u_k = \begin{cases} (-\hat{f}(\hat{x}_k) + K\hat{x}_k)/\hat{g}(\hat{x}_k), & k_i < k < k_{i+1}, \\ (-\hat{f}(x_k) + Kx_k)/\hat{g}(x_k), & k = k_i. \end{cases} \tag{19}$$

The control law using the approximated dynamics from (18) is given by

$$u_k = \begin{cases} (-\hat{W}_{f,k}^T \varphi_f(\hat{\bar{x}}_k) + K\hat{x}_k) \big/ \hat{W}_{g,k}^T \varphi_g(\hat{\bar{x}}_k), & k_i < k < k_{i+1}, \\ (-\hat{W}_{f,k}^T \varphi_f(\bar{x}_k) + Kx_k) \big/ \hat{W}_{g,k}^T \varphi_g(\bar{x}_k), & k = k_i. \end{cases} \tag{20}$$

To ensure the control law (20) is well-defined, i.e., $\hat{W}_{g_k}^T \varphi_g(\hat{\bar{x}}) \neq 0$ at all-time instants k, the estimate $\hat{g}(\hat{x}_k)$ is defined as

$$\hat{g}(\hat{x}_k) = \begin{cases} \hat{W}_{g,k}^T \varphi_g(\hat{\bar{x}}_k), \ \hat{W}_{g,k}^T \varphi_g(\hat{\bar{x}}) \geq g_{\min} \\ \hat{W}_{g,k-1}^T \varphi_g(\hat{\bar{x}}_{k-1}), \text{ otherwise} \end{cases}. \tag{21}$$

The augmented function approximation error can be written from (16) and (17) as

$$\tilde{\bar{F}}(x_k) = \bar{F}(x_k) - \hat{\bar{F}}(x_k) = \tilde{W}_k^T \Phi(\bar{x}_k) + \Xi(x_k), \ k_i \leq k < k_{i+1} \tag{22}$$

where $\tilde{\bar{F}} = [\tilde{f} \quad \tilde{g}]$ with $\tilde{f}(\bullet) = f(\bullet) - \hat{f}(\bullet)$ and $\tilde{g}(\bullet) = g(\bullet) - \hat{g}(\bullet)$ are the function approximation errors for f and g, respectively. The NN weight estimation error is denoted as $\tilde{W}_k = W - \hat{W}_k$.

3.2 EVENT TRIGGER ERROR DYNAMICS AND APERIODIC UPDATE LAW

The dynamics of the event-trigger error (12) using (3) and (15) for $k_i < k < k_{i+1}$ can be can be written as

$$e_{k+1}^s = x_{k+1} - \hat{x}_{k+1} = Ax_k + B\bar{F}(x_k)\bar{u}_k - A\hat{x}_k - B\hat{\bar{F}}(\hat{x}_k)\bar{u}_k$$

$$= Ae_k^s + B\tilde{\bar{F}}(x_k)\bar{u}_k + B(\hat{\bar{F}}(x_k) - \hat{\bar{F}}(\hat{x}_k))\bar{u}_k, \ k_i < k < k_{i+1}. \tag{23}$$

Recalling the event-based function approximation (17) and the augmented function approximation error (22), equation (23) can be expressed as

$$e_{k+1}^s = Ae_k^s + B\tilde{W}_k^T \Phi(\bar{x}_k)\bar{u}_k + B\hat{W}_k^T \left(\Phi(\bar{x}_k) - \Phi(\hat{\bar{x}}_k)\right)\bar{u}_k + B\Xi(x_k)\bar{u}_k$$

$$= Ae_k^s + B\tilde{W}_k^T\Phi(\bar{\mathbf{x}}_k)\bar{u}_k + B\Xi_k\bar{u}_k + B\hat{W}_k^T\tilde{\Theta}\left(\bar{x}_k,\hat{\bar{x}}_k\right)\bar{u}_k,\ k_i < k < k_{i+1}, \tag{24}$$

where $\tilde{\Theta}(\bar{x}_k,\hat{\bar{x}}_k) = \Phi(\bar{x}_k) - \Phi(\hat{\bar{x}}_k)$ and $\Xi_k \equiv \Xi(x_k)$ for brevity. Similarly, the event-trigger error dynamics at the trigger instants using (3) and (18) become

$$e_{k+1}^s = B\tilde{W}_k^T\Phi(\bar{x}_k)\bar{u}_k + B\Xi_k\bar{u}_k,\ k = k_i. \tag{25}$$

To ensure the convergence of the NN weight estimation error, $\tilde{W}_k$, the NN weight update law in an event-triggered context is selected as

$$\hat{W}_k = \hat{W}_{k-1} + \frac{\gamma_k\alpha\Phi(\bar{x}_{k-1})\bar{u}_{k-1}e_k^{s^T}B}{1+\left\|\Phi(\bar{x}_{k-1})\right\|^2\left\|\bar{u}_{k-1}\right\|^2} - \kappa\hat{W}_{k-1},\ k_{i-1} \le k < k_i, \tag{26}$$

where $\alpha > 0$ is the learning rate and $\kappa > 0$ is sigma modification term similar to that in traditional adaptive control [19]. The indicator function, γ_k, is defined as

$$\gamma_k = \begin{cases} 0, \text{ event is not triggered, } k_{i-1} < k < k_i, \\ 1, \text{ event is triggered, } k = k_i. \end{cases} \tag{27}$$

The indicator function enables the NN weights to be updated once an event is triggered, i.e., $\gamma_k = 1$. The event-trigger error e_k^s is first used to update the NN weights in (26) and then reset to zero for the next trigger. As the trigger instants are aperiodic in nature, the NN weights are updated in a non-periodic manner, as proposed. This saves the computation when compared to the traditional NN based control approaches [20].

The update law (26) needs both x_{k_i-1} and x_{k_i} at the trigger instant k_i for updating the NN weights and to reset the model state. As proposed, both the current and previous state vectors are transmitted as a single packet at the trigger instants.

The NN weight estimation error dynamics using (26) and forwarding one time step ahead, can be derived as

$$\tilde{W}_{k+1} = \tilde{W}_k - \frac{\gamma_k \alpha \Phi(\bar{x}_k)\bar{u}_k e_{k+1}^{s^T} B}{1+\left\|\Phi(\bar{x}_k)\right\|^2 \left\|\bar{u}_k\right\|^2} + \kappa \hat{W}_k \,,\ k_i \le k < k_{i+1} \,. \tag{28}$$

The convergence of the NN weight estimation $\tilde{W}_k$ requires the vector $\Phi(\bar{x}_k)\bar{u}_k$ in (28) satisfy the persistency of the excitation condition (PE) which is a well-known fact in traditional adaptive and NN based control [19]-[20], [23]-[24]. For completeness the definition of the PE condition is given below.

Definition 1[22]: *(Persistency of excitation)* A vector $\phi(x) \in \Re^n$ is said to be persistently exciting over an interval if there exist positive constants δ , c , d and $k_d \ge 1$, such that

$$cI \le \sum_{k=k_d}^{k+\delta} \phi(x_k)\phi^T(x_k) \le dI \,, \tag{29}$$

where I is the identity matrix of appropriate dimension.

Remark 2: A PE like condition for $\Phi(\bar{x}_k)\bar{u}_k$ can be achieved by adding an exploration noise to the control input [25]. This keeps the control input and, in turn, the system states away from zero. Further, the activation function also satisfies PE and $0 < \Phi_{\min} \le \left\|\Phi(x_k)\right\| \le \Phi_{\max}$ holds.

Lemma 1: Consider the adaptive SE (18) and the control law (20). Suppose the Assumptions 1 and 2 hold, the NN weights be initialized in a compact set and tuned by using (26), and the vector $\Phi(\bar{x}_k)\bar{u}_k$ satisfies the PE condition. Let k_0 be the initial trigger instant, k_p be p^{th} trigger instant for an integer p and $N \ge k_p$ is an integer representing the time instant. Then, the NN weight estimation error $\tilde{W}_k$ is bounded for all time and will converge to the ultimate bound when $k_i > k_p$ or, alternatively, for all time instants $k > k_0 + N$ provided the learning gains satisfy $0 < \alpha < 1/4$ and $0 < \kappa < 1/4$.

Proof: Refer to Appendix.

Note that the ultimate bound can be made arbitrarily small by selecting the proper design parameters and number of neurons as discussed in Remark 5. Next, the main results are claimed.

4. EVENT-TRIGGERING CONDITION AND STABILITY

In this section, the ultimate boundedness (UB) [20] of the closed-loop ETC system state vector and NN weight estimation error is presented by designing a suitable adaptive event-triggering condition.

4.1 CLOSED-LOOP SYSTEM DYNAMICS

The closed-loop dynamics of the ETC system can be derived by using (2) and (19). Consider the inter-event times, i.e., $k_i < k < k_{i+1}$, $\forall i \in \mathbb{N}$. The closed-loop dynamics can be written as

$$\begin{aligned} x_{k+1} &= Ax_k + Bf(x_k) + (Bg(x_k)(-\hat{f}(\hat{x}_k) + K\hat{x}_k)/\hat{g}(\hat{x}_k)) \\ &= A_c x_k + BKe_k^s + B\tilde{\bar{F}}(x_k)\bar{u}_k + B(\hat{\bar{F}}(x_k) - \hat{\bar{F}}(\hat{x}_k))\bar{u}_k. \end{aligned}$$

By using the NN estimation (17) , (18) and the function approximation error (22), the closed-loop dynamics become

$$x_{k+1} = A_c x_k + BKe_k^s + B\tilde{W}_k^T \Phi(\bar{x}_k)\bar{u}_k + B\Xi_k \bar{u}_k + B\hat{W}_k^T \tilde{\Theta}(\bar{x}_k, \hat{\bar{x}}_k)\bar{u}_k, \quad k_i < k < k_{i+1}. \tag{30}$$

Similarly, at the trigger instants, $k = k_i$, $\forall i \in \mathbb{N}$, the closed-loop dynamics using (2), (3), (19) and (20) can be written as

$$x_{k+1} = A_c x_k + B\tilde{W}_k^T \Phi(\bar{x}_k)\bar{u}_k + B\Xi_k \bar{u}_k. \tag{31}$$

The closed-loop dynamics of the SE can be derived by using (14) and (19) as

$$\hat{x}_{k+1} = \begin{cases} A_c \hat{x}_k, & k_i < k < k_{i+1}, \\ A_c x_k, & k = k_i. \end{cases} \tag{32}$$

The flowchart in Figure 3 shows the implementation of the adaptive MBETC scheme designed in Section 4.

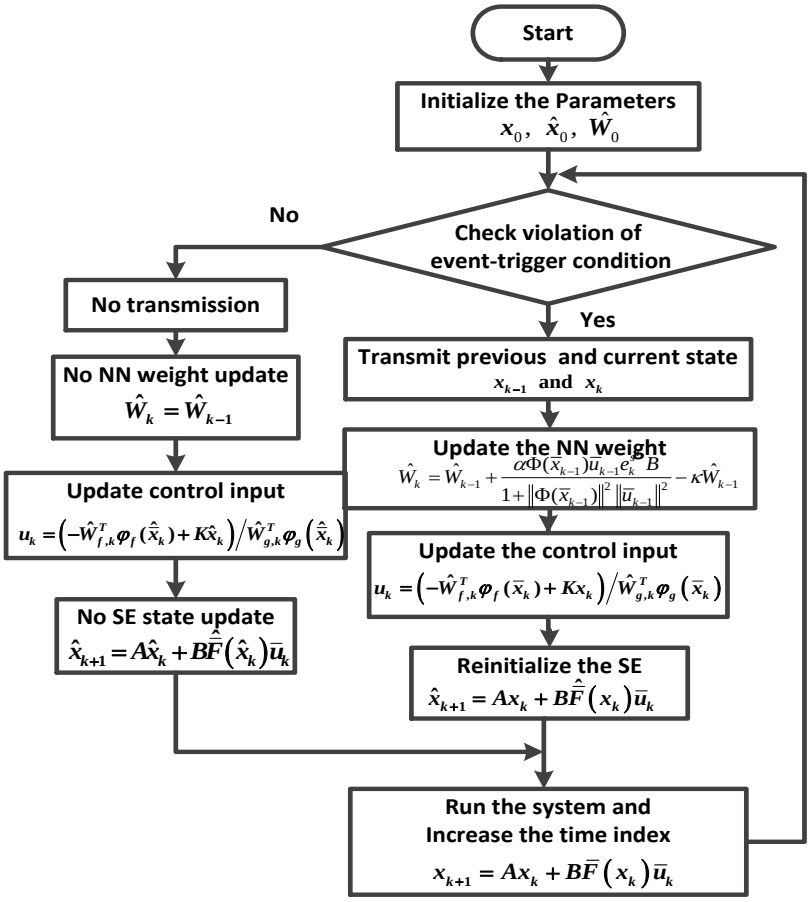

Figure 3. Flowchart of the proposed event-triggered control system.

4.2 MAIN RESULTS

In this section, we claim the main results by designing an adaptive event-triggering condition. The closed-loop stability of the adaptive MBETC is shown by evaluating a single Lyapunov function for both during the trigger instants and inter-event times. It is shown in [10], [17] that the Lyapunov function need not monotonically decrease both during the inter-event and event times [10]. Due to the aperiodic NN weight update, it is shown that the Lyapunov function may increase during the inter-event times but remains bounded. It is further shown that the bound during the inter-event times converges to the ultimate value with trigger of events. This is illustrated in Figure 4.

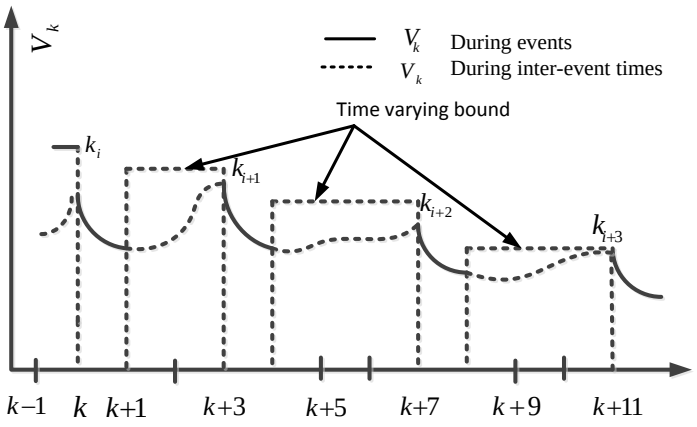

Figure 4. Evolution of the Lyapunov function during event and inter-event times

Consider the event-trigger error (12). We introduce the following condition as the event-triggering condition and given by

$$D\left(\left\|e_k^s\right\|\right) \le \mu_k^{ET} \left\|x_k\right\|, \tag{33}$$

where $\mu_k^{ET} = \sqrt{\Gamma \sigma_{\min}(Q) \Big/ (8\|K\|^2 \|\Lambda_1\| + 8L_\Phi^2 \|\bar{u}_k\|^2 \|\hat{W}_k\|^2 \|\Lambda_1\|)}$ is the threshold coefficient and $0 < \Gamma < 1$. The matrices Λ_1 and Q are positive definite matrices and satisfy the Lyapunov equation $\bar{A}_c^T \Lambda_1 \bar{A}_c - \Lambda_1 = -Q$ with $\bar{A}_c = \sqrt{2} A_c$ as in Remark A.1. The minimum singular value of Q is denoted as $\sigma_{\min}(Q)$. The dead zone operator $D(\bullet)$ is defined by

$$D\left(\left\|e_k^s\right\|\right) = \begin{cases} \left\|e_k^s\right\|, & \text{if } \left\|x_k\right\| > B_x, \\ 0, & \text{otherwise,} \end{cases} \tag{34}$$

with B_x the desired ultimate bound for the system state vector. The events are triggered at the violation of the inequality (33) .

Remark 3: The threshold coefficient μ_k^{ET} in (33) is a function of the NN weight estimate, $\left\|\hat{W}_k\right\|$ and gets updated by (26). Therefore, the event-triggering condition (33) becomes adaptive. This helps in generating the required number of events for the function approximation during the learning phase of the NNs as discussed in Section 2.2.1. Further, μ_k^{ET} is also function of the control gain K. The choice of control gain K is based on the desired closed-loop performance and stability of the system such that $\bar{A}_c$ is Schur. This implies, the event-triggering condition is also driven by the system performance. Hence, for different choice of K, the triggering condition will ensure the required number of events to achieve the desired performance.

Remark 4: The dead-zone operator (34) is utilized in the event-triggering condition in order to reset the event-trigger error e_k^s to zero once the state vector is within the ultimate bound. This avoids unnecessary triggering of events due to the NN reconstruction error of the functions.

Theorem 1: Consider the nonlinear discrete time system (1) along with the NN-based SE given in (18). Assume Assumptions 1 through 3 hold and the NN initial weight matrix $\hat{W}_0$ be initialized in a compact set**.** Suppose the system state vectors, x_{k_i} and x_{k_i-1}, are transmitted, the SE state vector, $\hat{x}_k$, is reinitialized and the NN weights are updated using (26) at the violation of the inequality (33). Let k_0 be the initial trigger instant, k_p be p^{th} trigger instant for any positive integer p and $N \geq k_p$ is an integer represents the time instant. Then, the control input (20) ensures the closed-loop event-triggered system state vector, x_k, the SE state vector, $\hat{x}_k$, and the NN weight estimation error, $\tilde{W}_k$, are bounded

for all time and converge to the ultimate bound for all trigger instants $k_i > k_p$ or, alternatively, for all $k > k_0 + N$ provided learning gains satisfy $0 < \alpha < 1/4$ and $0 < \kappa < 1/4$.

Proof: Refer to the Appendix.

Remark 5: From the proof of Theorem 1 (see Appendix), the bounds on the system state vector, B_x, and NN weight estimation errors, $B_{\tilde{W}}$, depend upon the traditional NN reconstruction error, $\Xi_{\max}$, and the design parameters, α and κ. Through proper selection of the number of neurons in the hidden layer, and the design parameters α and κ, the bounds B_x and $B_{\tilde{W}}$ can be made arbitrary small (see simulation results).

The minimum inter-event time, $\delta k_{\min} = \min_{i \in \mathbb{N}} (\delta k_i)$, where $\delta k_i = k_{i+1} - k_i$ for $i \in \mathbb{N}$, implicitly defined by the event-triggering condition (33), is the minimum time required for the event-trigger error to evolve from zero and reach the event-triggering threshold over all inter-event times. In the case of a discrete-time system, which can be considered as discretized version of a continuous time system with a suitable fixed sampling time, trivially the minimum inter-event time is the sampling time [4]. Further, in our case of model-based adaptive NN ETC, minimum inter-inter time may be one sampling time during the learning phase but the inter-event times increases with the convergence of NN weight estimation error and thereby reducing the transmission.

5. SIMULATION RESULTS

In this section, the proposed NN-based MBETC scheme is evaluated by using two examples.

5.1 EXAMPLE 1

A second order SISO nonlinear discrete time system was selected for simulations whose dynamics are given as

$$\begin{aligned} x_{1,k+1} &= x_{2,k}, \\ x_{2,k+1} &= f(x_k) + g(x_k)u_k, \end{aligned} \tag{35}$$

where $f(x_k) = x_{2,k} / 1 + x_{1,k}^2$ and $g(x_k) = 2 + \sin(x_{1,k})$.

The following parameters were considered during the simulations. The initial states of the system and the SE were selected to be $[3 \quad 2]^T$ since first event is considered at k_0. Initial NN weights $V \in \Re^{2\times 15}$, $\hat{W}_{f,0} \in \Re^{15}$ and $\hat{W}_{g,0} \in \Re^{15}$ were chosen randomly from a uniform distribution in the interval [0, 1] with 15 neurons each in the hidden layers. The activation functions used were symmetric sigmoid functions ($tanh(\bullet)$) for both the NNs with learning gains $\alpha = 0.24$ and $\kappa = 10^{-5}$. The control gain $K = [0.35 \quad 0.2]$ such that the matrix $\bar{A}_c = \sqrt{2}A_c$ is Schur. The event-triggering condition was derived from (33) with $g_{\min} = 1$ and $\Gamma = .99$. The Lipschitz constant L_Φ was computed as $L\|V\| = 3.28$ with $L = 1$. The system was simulated for 15 sec. with a sampling time of 0.01 sec, i.e., 1500 sampling instants. The UB for the system state vector was chosen to be 10^{-3}.The simulation results are presented in Figure 5, 6 and 7.

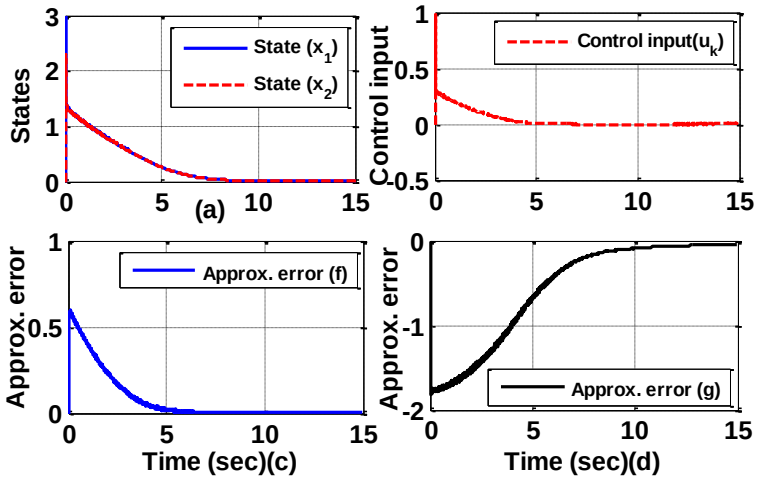

Figure 5. Convergence of (a) state vector, (b) control input, (c) function approximation error $\tilde{f}(x_k)$, and (d) function approximation error $\tilde{g}(x_k)$.

Figure 5 (a) shows the convergence of the system state vector close to zero with the event-based approximated control input in Figure 5 (b). The NN approximation errors of the nonlinear functions $f(x_k)$ and $g(x_k)$ are shown in Figure 5 (c) and (d), respectively. Due to NN initial weights being far away from the target, large initial errors are noticed in the plot and finally they converge to a bound close to zero. The boundedness of these errors close to zero validated the event-based approximation discussed in Section 4.

Next, the performance in terms of the triggering of events is plotted in Figure 6 and 7. Figure 6 (a) shows the evolution of the state dependent event-trigger threshold and the error. The event-trigger error (see the zoomed figure in Figure 6 (a)) resets to zero once the error reaches the threshold and the system states were transmitted. Figure 6 (b) illustrates the count on the number of trigger instants that have occurred with respect to the total number of sampling instants. It was found that a total of 306 events occurred out of 1500 sampling instants. In addition, the plot indicates that the events are triggered

frequently at the initial phase as a result of large approximation error resulting from random initialization of NN weights. As the NN weights are updated and converge close to the target weights, inter-trigger times increase. As expected, changing initial NN weights resulted in different number of events for the convergence of the weights.

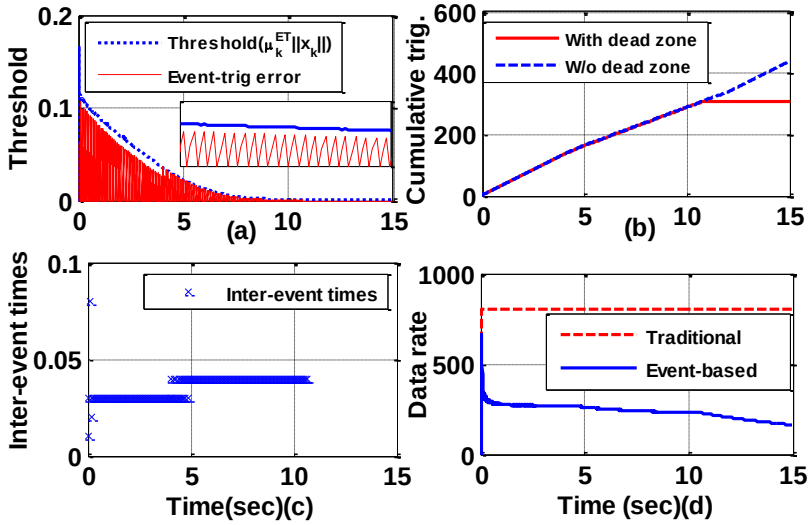

Figure 6. Performance of the model-based adaptive NN ETC: (a) evolution of event-trigger threshold and event-trigger error, (b) cumulative number of trigger instants with and without dead zone operator, (c) inter-event time, and (d) comparison of the data rate between the traditional periodic transmission and event-triggered transmission.

The reduction in the number of cumulative events (y-axis in Figure 6 (b)) demonstrates the effectiveness of the event-triggered scheme in reducing the number of state vector transmissions over the network in comparison to a traditional periodic sampled discrete time control. The durations between two consecutive transmissions are shown in Figure 6(c) and are aperiodic in nature. Assuming every packet size of 8 bit data, a comparison plot for the data rate in bits/sec is shown in Figure 6 (d). In the case of traditional discrete time system, the data rate is constant, i.e., 800 bits/sec. In contrast, in

the proposed ETC, the data rate reduces over time since the transmissions are reduced and finally reaches to 100 bits/sec. This confirms a reduction in bandwidth usage and proves the effectiveness of the approach. Further, the NN weights are updated 306 times thus reducing the computation for approximating the unknown nonlinear functions when compared to traditional NN based approach. However, the use of mirror adaptive SE for evaluation of the event-triggering condition requires additional computation.

A comparison between the trigger mechanisms with and without a dead zone operator, in terms of cumulative number of event-trigger instants, is presented in Figure 6 (b). When the dead zone operator is not used, as shown in Figure 6(b) (dotted line), the events trigger continuously due to the NN reconstruction error even the system state vector is inside the ultimate bound. Hence, the dead-zone operator is necessary to reset the event-trigger error to zero once the state vector converge and stay inside the ultimate bound. This stops the unnecessary triggering of events as shown in Figure 6 (b) (bold line).

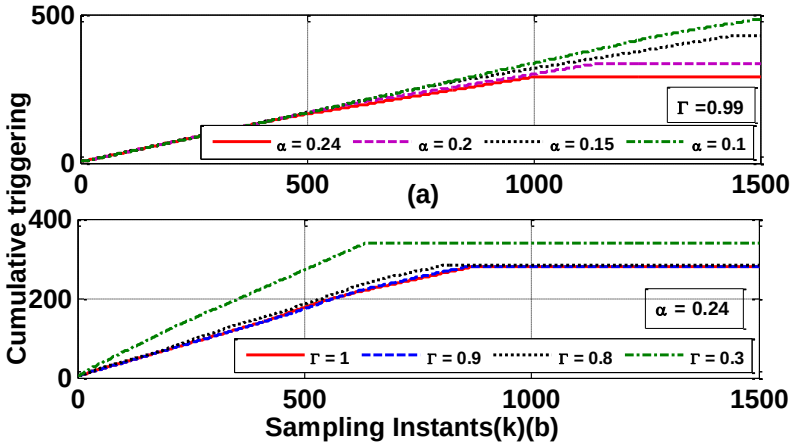

Figure 7. Cumulative number of events with different values of the learning gain α and event-trigger parameter Γ.

Further, the effect of different learning gains α and event-trigger parameters Γ on the number of events is shown in Figure 7. As shown in Figure 7 (a), for different values of α the cumulative number of events is different. This is due to the change in convergence rate of the NN weight updates. The number of cumulative triggers reduced with an increase in value Γ since the threshold value increases with an increase in Γ. Note that Lyapunov stability is a sufficient condition. Therefore, the event-trigger threshold for $\Gamma = 1$ still maintains the stability of the system.

5.2 EXAMPLE 2

In this example, another second order system as in (35) was chosen where the system dynamics are given by

$$f(x_k) = x_{1,k}^2 x_{2,k} / (1 + x_{1,k}^2 + x_{2,k}^2) \text{ and } g(x_k) = 1 + \left(2/(1 + x_{1,k}^2 + x_{2,k}^2)\right).$$

The simulation parameters were as follows. The initial vales for the system and SE states were $[1.5 \quad 2.5]^T$. The initial NN weights, $V \in \Re^{2\times 16}$, $\hat{W}_{f,0} \in \Re^{16}$ and $\hat{W}_{g,0} \in \Re^{16}$ were chosen randomly in the interval [0, 1] with 16 neurons each in the hidden layers. Symmetric sigmoid functions were used as activation functions for both the NNs. Design parameters were selected as $\alpha = 0.24$, $\kappa = 10^{-5}$, $g_{\min} = 1$, $\Gamma = .99$, $L_\Phi = 3.6$ and $K = [0.3 \quad 0.25]$. The system was simulated for 5 sec. with a sampling time of 0.01 sec, i.e., 500 sampling instants. The ultimate bound threshold of system state vector was chosen to be 8×10^{-3}.

The convergence of the system state and the control input are shown in Figure 8 (a) and (b), respectively. The NN approximation errors $\tilde{f}$ and $\tilde{g}$ are illustrated in Figure

8 (c) and (d), respectively. The cumulative number of triggers was observed to be 80 out of 500 sampling instants implying the saving in network resources and computation.

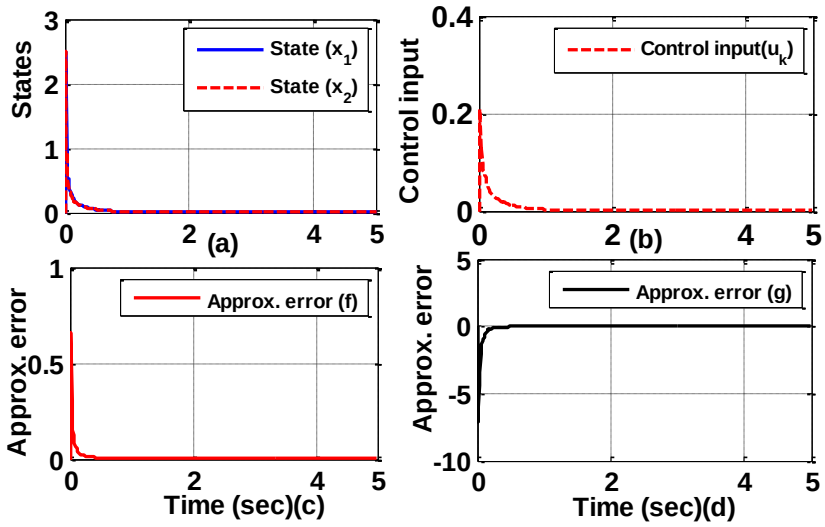

Figure 8. Convergence of (a) state vector, (b) control input, (c) function approximation error $\tilde{f}(x_k)$, and (d) function approximation error $\tilde{g}(x_k)$.

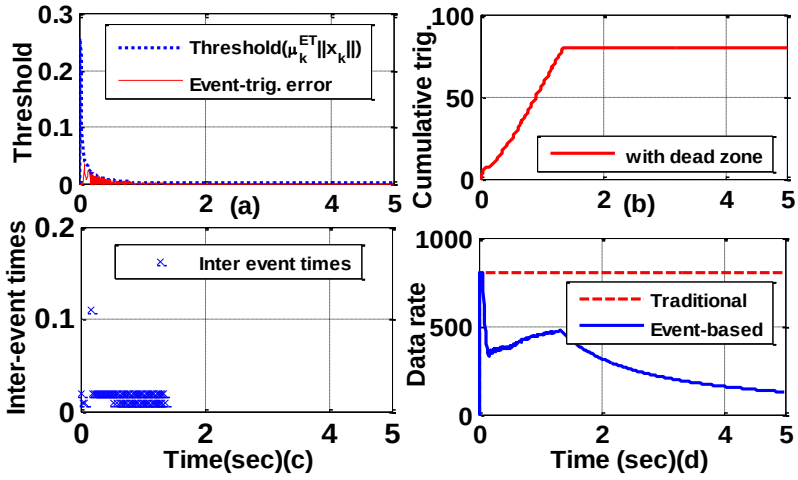

Figure 9. Performance of the model-based adaptive NN ETC: (a) evolution of event-trigger threshold and event-trigger error, (b) cumulative number of trigger instants vs the total number of sampling instants with and without dead zone operator, (c) inter-event time, and (d) comparison of the data rate between the periodic transmission and event-triggered transmission.

From both the examples, it is clear that the adaptive triggering condition is able to generate required number of triggers for event-based function approximation with aperiodic update law. Further, the reduction in the number of transmission verified the saving in communication bandwidth.

6. CONCLUSIONS

In this paper, a NN based adaptive ETC scheme for an uncertain nonlinear discrete-time system was introduced. Approximation of system dynamics by using NN was accomplished in the context of reduced event sampled communication. Two linearly parameterized NNs approximate the unknown nonlinear functions quite satisfactorily. The novel adaptive event-triggering condition ensured the stability and desired performance of the complete uncertain system. In addition, the simulation results proved the efficacy of the proposed algorithm in terms of reducing the network traffic. It was observed that the number of triggered instants vary with initial NN weights and learning gain. Though a stabilizing controller was designed, it is not optimal. Hence, the design of event-based optimal controller for uncertain systems will be as part of future research.

7. REFERENCES

[1] K. Astrom and B. Bernhardsson, "Comparison of Riemann and Lebesgue sampling for first order stochastic systems," in *Proceedings of the 42st IEEE Conference on Decision and Control*, vol. 2, Las Vegas, Nevada, USA, Dec. 2002, pp. 2011–2016.

[2] A. Anta, and P. Tabuada, "To sample or not to sample: self-triggered control for nonlinear system," *IEEE Transactions on Automatic Control*, vol. 55, no. 9, pp. 2030-2042, Sep. 2010.

[3] 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.

[4] A. Eqtami, D. V. Dimarogonas, and K. J. Kyriakopoulos, "Event-triggered control for discrete-time systems," in *Proceedings of the American Control Conference*, Baltimore, MD, USA, Jun. 2010, pp. 4719-4724.

[5] M. Donkers and W. Heemels, "Output-based event-triggered control with guaranteed $\mathcal{L}_\infty$ -gain and improved and decentralised event-triggering," *IEEE Transactions on Automatic Control*, vol. 57, no. 6,, pp. 1362–1376, Jun. 2012.

[6] J. Lunze and D. Lehmann, "A state-feedback approach to event-based control," *Automatica*, vol. 46, no. 1, pp. 211–215, Jan. 2010.

[7] C. Stocker and J. Lunze, "Event-based control of nonlinear systems: An input-output linearization approach," in *Proceedings of the 50th IEEE Conference on Decision and Control and European Control Conference*, Orlando, FL, USA, Dec. 2011, pp. 2541–2546.

[8] E. Garcia and P. J. Antsaklis, "Model-based event-triggered control for systems with quantization and time-varying network delays," *IEEE Transactions on Automatic Control*, vol. 58, no. 2, pp. 422-434, Feb. 2013.

[9] E. Garcia and P. J. Antsaklis, "Parameter estimation in time-triggered and event-triggered model-based control of uncertain system," *International Journal of Control*, vol. 85, no. 9, pp. 1327-1342, Apr. 2012.

[10] X. Wang and M. D. Lemmon, "On event design in event-triggered feedback systems," *Automatica*, vol. 47, no. 10, pp. 2319-2322, Oct. 2011.

[11] W. Heemels and M. Donkers, "Model-based periodic event-triggered control for linear systems," *Automatica*, vol. 49, no. 3, pp. 698-711, Mar.. 2013.

[12] X. Wang and N. Hovakimyan, "L1 adaptive control of event-triggered networked systems," in *Proceedings of the American Control Conference*, Baltimore, MD, USA, Jun. 2010, pp. 2458-2463.

[13] P. Tallapragada and N. Chopra, "On event triggered trajectory tracking for control affine nonlinear systems," in *Proceedings of the 50th IEEE Conference on Decision and Control and European Control Conference*, Orlando, FL, USA, Dec. 2011, pp. 5377–5382.

[14] A. Sahoo, Hao Xu, and S. Jagannathan, "Adaptive event-triggered control of an uncertain linear discrete time system using measured input-output data," in *Proceedings pf the American Control Conference*, Washington, DC, USA, Jun. 2013, pp. 5692-5697.

[15] A. Sahoo, Hao Xu, and S. Jagannathan, "Neural network-based adaptive event-triggered control of affine nonlinear discrete time systems with unknown internal dynamics," in *Proceedings of the American Control Conference*, Washington, DC, USA, Jun. 2013, pp. 6433-6438.

[16] A. Sahoo, Hao Xu, and S. Jagannathan, "Near optimal event-based control of nonlinear discrete time system in affine form with measured input output data," *in Proceedings of the International Joint Conference on Neural Networks, Beijing, China, Jul.* 2014, pp. 3671-3676.

[17] H. Ye, A. Michel, and L. Hou, "Stability theory for hybrid dynamical systems," *IEEE Transactions on Automatic Control*, vol. 43, no. 4, pp. 461–474, Apr .1998.

[18] A. Yesildirek and F. Lewis, "Feedback linearization using neural networks," *Automatica*, vol. 31, no. 11, pp. 1659-1664, Nov. 1995.

[19] K. S. Narendra and A. M. Annaswamy, *Stable Adaptive Systems*. New York: Dover Publication, 2005.

[20] S. Jagannathan, *Neural Network Control of Nonlinear Discrete-time Systems.* Boca Raton, FL: CRC Press, 2006.

[21] H. Khalil, *Nonlinear Systems*, 3rd ed. New Jersey: Prentice Hall, 2002.

[22] Michael Green and John B. Moore, "persistency of excitation in linear systems," *Systems & Control Letters*, vol. 7, pp. 351 -360, Sep. 1986.

[23] G. C. Goodwin and K. S. Sin, *Adaptive Filtering Prediction and Control*. New York: Dover Publication, 2009.

[24] D. Gorinevsky, "On the persistency of excitation in radial basis function network identification of nonlinear system," *IEEE Transactions on Neural Networks*, vol. 6, no. 5, pp 1237-1244, Sep. 1995.

[25] H. Xu, S. Jagannathan, and F. Lewis, "Stochastic optimal control of unknown linear networked control system in the presence of random delays and packet losses," *Automatica*, vol. 48, no. 6, pp. 1017-1030, Jun. 2012.

APPENDIX

Proof of Lemma 1: The NN weights are updated only at the trigger instants and held during the inter-events times. Thus, the proof for the ultimate boundedness of the NN weight estimation error is carried out by evaluating a Lyapunov function candidate for both the cases as follows.

Case I: At the event-triggered instants ($k = k_i, \forall i \in \mathbb{N}$)

Consider the Lyapunov function given by

$$V_{\tilde{W},k} = tr\{\tilde{W}_k^T \tilde{W}_k\} \ . \tag{A.1}$$

The first difference, $\Delta V_{\tilde{W},k} = tr\{\tilde{W}_{k+1}^T \tilde{W}_{k+1}\} - tr\{\tilde{W}_k^T \tilde{W}_k\}$, along the weight estimation error dynamics (28) with the indicator function $\gamma_k = 1$ can be written as

$$\Delta V_{\tilde{W},k} = tr\left\{\left(\tilde{W}_k - \frac{\alpha\Phi(\bar{x}_k)\bar{u}_k}{1+\left\|\Phi(\bar{x}_k)\right\|^2\left\|\bar{u}_k\right\|^2} e_{k+1}^{s^T} B + \kappa\hat{W}_k\right)^T \left(\tilde{W}_k - \frac{\alpha\Phi(\bar{x}_k)\bar{u}_k}{1+\left\|\Phi(\bar{x}_k)\right\|^2\left\|\bar{u}_k\right\|^2} e_{k+1}^{s^T} B + \kappa\hat{W}_k\right)\right\}$$
$$-tr\{\tilde{W}_k^T \tilde{W}_k\}.$$

Substitute the error dynamics in (25). Applying Cauchy-Schwartz (C-S) inequality with definitions $\hat{W}_k = W - \tilde{W}_k$ and $B^T B = 1$, the first difference satisfies

$$\Delta V_{\tilde{W},k} \le -\frac{\alpha}{1+\left\|\Phi(\bar{x}_k)\right\|^2\left\|\bar{u}_k\right\|^2} tr\left\{\tilde{W}_k^T \Phi(\bar{x}_k)\bar{u}_k\bar{u}_k^T\Phi^T(\bar{x}_k)\tilde{W}_k\right\} + \frac{\alpha}{1+\left\|\Phi(\bar{x}_k)\right\|^2\left\|\bar{u}_k\right\|^2} tr\left\{\Xi_k\bar{u}_k\bar{u}_k^T\Xi_k^T\right\}$$
$$+4\alpha^2\Psi\frac{\bar{u}_k^T\Phi^T(\bar{x}_k)\Phi(\bar{x}_k)\bar{u}_k}{(1+\left\|\Phi(\bar{x}_k)\right\|^2\left\|\bar{u}_k\right\|^2)^2} tr\{(\tilde{W}_k^T\Phi(\bar{x}_k)\bar{u}_k)(\tilde{W}_k^T\Phi(\bar{x}_k)\bar{u}_k)^T\} - 2\kappa tr\{\tilde{W}_k^T\tilde{W}_k\} + 4\kappa^2 tr\{W^T W\}$$
$$+4\kappa^2 tr\{\tilde{W}_k^T\tilde{W}_k\} + 2\kappa tr\{\tilde{W}_k^T W\} + 4\alpha^2\frac{\bar{u}_k^T\Phi^T\left(\bar{x}_k\right)\Phi(\bar{x}_k)\bar{u}_k}{(1+\left\|\Phi(\bar{x}_k)\right\|^2\left\|\bar{u}_k\right\|^2)^2} tr\left\{(B\Xi_k\bar{u}_k)(B\Xi_k\bar{u}_k)^T\right\}$$

Observe that $\frac{\bar{u}_k^T\Phi^T(\bar{x}_k)\Phi(\bar{x}_k)\bar{u}_k}{1+\left\|\Phi(\bar{x}_k)\right\|^2\left\|\bar{u}_k\right\|^2} \le \frac{\left\|\Phi(\bar{x}_k)\right\|^2\left\|\bar{u}_k\right\|^2}{1+\left\|\Phi(\bar{x}_k)\right\|^2\left\|\bar{u}_k\right\|^2} \le 1$. Therefore, the first difference is upper bounded as

$$\Delta V_{\tilde{W},k} \le -\frac{\alpha(1-4\alpha)}{1+\left\|\Phi(\bar{x}_k)\right\|^2\left\|\bar{u}_k\right\|^2} tr\left\{\tilde{W}_k^T\Phi(\bar{x}_k)\bar{u}_k\bar{u}_k^T\Phi^T(\bar{x}_k)\tilde{W}_k\right\} + \frac{\alpha(1+4\alpha)}{1+\left\|\Phi(\bar{x}_k)\right\|^2\left\|\bar{u}_k\right\|^2}$$
$$\times tr\left\{\Xi_k\bar{u}_k\bar{u}_k^T\Xi_k^T\right\} + 2\kappa tr\left\{\tilde{W}_k^TW\right\} - 2\kappa tr\left\{\tilde{W}_k^T\tilde{W}_k\right\} + 4\kappa^2 tr\left\{W^TW\right\} + 4\kappa^2 tr\left\{\tilde{W}_k^T\tilde{W}_k\right\}.$$

By using the inequality $2tr\left\{A^TB\right\} \le \left\|A\right\|^2 + \left\|B\right\|^2$ and Frobenius norm, the first difference leads to

$$\Delta V_{\tilde{W},k} \le -\frac{\alpha(1-4\alpha)}{1+\left\|\Phi(\bar{x}_k)\right\|^2\left\|\bar{u}_k\right\|^2}\left\|\tilde{W}_k^T\Phi(\bar{x}_k)\bar{u}_k\right\|^2 + \kappa W_{\max}^2 + \kappa\left\|\tilde{W}_k\right\|^2 - 2\kappa\left\|\tilde{W}_k\right\|^2 + 4\kappa^2W_{\max}^2 + 4\kappa^2\left\|\tilde{W}_k\right\|^2$$
$$+\left(\alpha(1+4\alpha)\left\|\bar{u}_k\right\|^2\Xi_{\max}^2 \big/ (1+\left\|\Phi(\bar{x}_k)\right\|^2\left\|\bar{u}_k\right\|^2)\right).$$

Since $0 < \Phi_{\min} \le \left\|\Phi\left(\bar{x}_k\right)\right\|$ is ensured due the PE condition as discussed in Remark 2, the following inequality holds. $\dfrac{\left\|\bar{u}_k\right\|^2}{1+\left\|\Phi\left(\bar{x}_k\right)\right\|^2\left\|\bar{u}_k\right\|^2} = \dfrac{\left\|\Phi\left(\bar{x}_k\right)\right\|^2\left\|\bar{u}_k\right\|^2}{(1+\left\|\Phi\left(\bar{x}_k\right)\right\|^2\left\|\bar{u}_k\right\|^2)\left\|\Phi\left(\bar{x}_k\right)\right\|^2} \le \dfrac{1}{\Phi_{\min}^2}$. The first difference using the above inequality leads to

$$\Delta V_{\tilde{W},k} \le -\left(\alpha(1-4\alpha)\big/(1+\left\|\Phi(\bar{x}_k)\right\|^2\left\|\bar{u}_k\right\|^2)\right)\left\|\tilde{W}_k^T\Phi(\bar{x}_k)\bar{u}_k\right\|^2 - \kappa(1-4\kappa)\left\|\tilde{W}_k\right\|^2 + B_M^{\tilde{W}},$$

where $B_M^{\tilde{W}} = (\alpha(1+4\alpha)\Xi_{\max}^2/\Phi_{\min}^2) + (\kappa+4\kappa^2)W_{\max}^2$ and $0 < \alpha < 1/4$. Dropping the first negative term, it holds that

$$\Delta V_{\tilde{W},k} \le -\beta\left\|\tilde{W}_k\right\|^2 + B_M^{\tilde{W}},\ k = k_i\ ,\ \forall i \in \mathbb{N}, \tag{A.2}$$

where $\beta = \kappa(1-4\kappa) > 0$ by selecting $0 < \kappa < 1/4$. From (A.2), the first difference of the Lyapunov function, $\Delta V_{\tilde{W},k}$, is less than zero as long as $\left\|\tilde{W}_k\right\|^2 > B_M^{\tilde{W}}/\beta \equiv B_{UB}^{\tilde{W}}$. Therefore by using Lyapunov theorem [20], the NN weight estimation error $\tilde{W}_k$ is bounded at the trigger instants provided the vector $\Phi(\bar{x}_k)\bar{u}_k$ satisfies the PE condition.

Case II: During the inter-event times ($k_i < k < k_{i+1}$, $\forall i \in \mathbb{N}$)

Consider the Lyapunov function in (A.1). Along the NN weight estimation error dynamics (28) with $\gamma_k = 0$, the first difference of $V_{\tilde{W},k}$ can be expressed as

$$\Delta V_{\tilde{W},k} = tr\{\tilde{W}_{k+1}^T \tilde{W}_{k+1}\} - tr\{\tilde{W}_k^T \tilde{W}_k\} = tr\{\tilde{W}_k^T \tilde{W}_k\} - tr\{\tilde{W}_k^T \tilde{W}_k\} = 0 . \tag{A.3}$$

From (A.3), the NN weight estimation error $\tilde{W}_k$ is a constant during the inter-event times. Since, the NN weights are bounded at the trigger instants as demonstrated in Case I, and the initial weights are being finite, the weight estimation error, $\tilde{W}_k$, is bounded during the inter-event times.

From both the cases, we need to show that the NN weight estimation error converges to the ultimate bound. The first difference (A.2) in Case I for $k = k_i$ can be expressed as

$$V_{\tilde{W},k_i+1} - V_{\tilde{W},k_i} = tr\left\{\tilde{W}_{k_i+1}^T \tilde{W}_{k_i+1}\right\} - tr\left\{\tilde{W}_{k_i}^T \tilde{W}_{k_i}\right\}$$

$$= \left\|\tilde{W}_{k_i+1}\right\|^2 - \left\|\tilde{W}_{k_i}\right\|^2 \leq -\beta \left\|\tilde{W}_{k_i}\right\|^2 + B_M^{\tilde{W}} .$$

Rearranging the above expression one can express the above inequality as

$$\left\|\tilde{W}_{k_i+1}\right\| \leq (1-\beta)\left\|\tilde{W}_{k_i}\right\|^2 + B_M^{\tilde{W}} . \tag{A.4}$$

It is clear that $0 < 1-\beta < 1$ by the choice of $0 < \kappa < 1/4$. Further, $\tilde{W}_k$ during the inter-event times, from (A.3) in Case II, remains constant. Thus, $\left\|\tilde{W}_{k_i}\right\| = \left\|\tilde{W}_{k_{i-1}+1}\right\|$ for $k_{i-1} < k < k_i$, $\forall i \in \mathbb{N}$. Therefore, (A.4) can be rewritten as

$$\left\|\tilde{W}_{k_i+1}\right\| \leq (1-\beta)\left\|\tilde{W}_{k_{i-1}+1}\right\|^2 + B_M^{\tilde{W}} . \tag{A.5}$$

Solving the difference inequality in (A.5) recursively, with initial NN weight estimation error $\left\|\tilde{W}_{k_0}\right\| = \left\|\tilde{W}_0\right\| = B_{\tilde{W}0}$, the NN weight estimation error in (A.5) can be expressed as

$$\left\|\tilde{W}_{k_i+1}\right\|^2 \leq (1-\beta)^{i+1} B_{W0}^2 + ((1-(1-\beta)^{i+1}/\beta) B_M^{\tilde{W}} \equiv B_i^{\tilde{W}}. \tag{A.6}$$

Therefore, the constant upper bound on the NN weight estimation error during the inter-event times from (A.6) is given by

$$\left\|\tilde{W}_k\right\|^2 \leq (1-\beta)^{i+1} B_{W0}^2 + ((1-(1-\beta)^{i+1}/\beta) B_M^{\tilde{W}} \equiv B_i^{\tilde{W}}, \tag{A.7}$$

for $k_i < k < k_{i+1}$, $\forall i \in \mathbb{N}$.

The NN weights are initialized with a finite value and the target weights are bounded. Therefore, the initial NN weight estimation error $\left\|\tilde{W}_{k_0}\right\| = \left\|\tilde{W}_0\right\| = B_{\tilde{W}0}$ is bounded. Further, from (A.6), $B_i^{\tilde{W}}$ is a converging sequence of piecewise constant functions since β satisfies $0 < \beta < 1$. Therefore, there exists an integer p (number of events) such that for the number of events $i > p$, the upper bound $B_i^{\tilde{W}}$ converges to the ultimate bound, i.e., $B_i^{\tilde{W}} \to B_{UB}^{\tilde{W}}$ for all event-trigger instants $k_i > k_p$ where $B_{UB}^{\tilde{W}} = B_M^{\tilde{W}}/\beta$ from (A.2).

Consequently, from Case I and Case II, the NN weight estimation error $\tilde{W}_k$ is bounded for all time instants and converges to the ultimate bound when $k_i > k_p$. Since k_i is a subsequence of k, the NN weight estimation error $\tilde{W}_k$ is UB for $k > k_0 + N$ where $N \geq k_p$ is a positive integer. ■

Proof of Theorem 1: The proof of the theorem is completed by considering two cases, i.e., at the event triggered instants and during the inter-event times. The first difference of the Lyapunov function is evaluated for both the cases and combined to show the UB.

Case I: Event triggered instants ($k = k_i, \forall i \in \mathbb{N}$)

Consider the Lyapunov function candidate given as

$$V_k = V_{x,k} + \eta V_{\hat{x},k} + \varpi V_{A,k} + \Psi V_{\tilde{W},k} + \Pi V_{B,k}, \tag{A.8}$$

where $V_{x,k} = x_k^T \Lambda_1 x_k$, $V_{\hat{x},k} = \hat{x}_k^T \Lambda_2 \hat{x}_k$, $V_{A,k} = (\hat{x}_k^T \Lambda_2 \hat{x}_k)^2$, $V_{\tilde{W},k} = tr\{\tilde{W}_k^T \tilde{W}_k\}$, $V_{B,k} = tr\{\tilde{W}_k^T \tilde{W}_k\}^2$. The matrices Λ_1 and Λ_2 are symmetric positive definite matrices and satisfy the Lyapunov equations $\bar{A}_c \Lambda_1 \bar{A}_c - \Lambda_1 = -Q$ and $A_c \Lambda_2 A_c - \Lambda_2, = -\bar{Q}$ where $\bar{A}_c = \sqrt{2} A_c$. The matrices Q and $\bar{Q}$ are positive definite matrices. The constant coefficients are defined as

$$\eta = \{17\|K\|^2 \Xi_{\max}^2 \|\Lambda_1\| / g_{\min}^2 \sigma_{\min}(\bar{Q}),\ 9\|\Lambda_1\| \Xi_{\max}^2 \|K\|^2 / g_{\min}^2 \sigma_{\min}(\bar{Q})\},$$

$$\varpi = \max\left\{9\|K\|^2 \|\Lambda_1\| \Phi_{\max}^2 / \{g_{\min}^2 \sigma_{\min}(\bar{\mathbf{Q}}) \sigma_{\min}(A_c^T \Lambda_2 A_c + \Lambda_2)\},\right.$$
$$\left. 5\|\Lambda_1\| \|K\|^4 / g_{\min}^2 \sigma_{\min}(\bar{Q}) \sigma_{\min}(A_c^T \Lambda_2 A_c + \Lambda_2)\right\},$$

$$\Psi = 2(4\Phi_{\max}^2 \|\Lambda_1\| g_{\min}^2 + 16\|\Lambda_1\| \Phi_{\max}^2 (\Phi_{\max}^2 W_{\max}^2 + \Xi_{\max}^2))/g_{\min}^2 \beta, \text{ and}$$

$$\Pi = 42\Phi_{\max}^4 \|\Lambda_1\| / \{g_{\min}^2 \beta(2-\beta)\}$$

with $\sigma_{\min}(\bullet)$ is the minimum singular value.

For brevity we will compute the first difference of each term in (A.8) individually and combine them at the final step to obtain the overall first difference. Consider the first term, $V_{x,k} = x_k^T \Lambda_1 x_k$. The first difference along the system dynamics (31) can be expressed as

$$\Delta V_{x,k} = \left[A_c x_k + B\tilde{W}_k^T \Phi(\bar{x}_k)\bar{u}_k + B\Xi_k \bar{u}_k\right]^T \Lambda_1 \left[A_c x_k + B\tilde{W}_k^T \Phi(\bar{x}_k)\bar{u}_k + B\Xi_k \bar{u}_k\right] - x_k^T \Lambda_1 x_k.$$

Applying C-S inequality, one can arrive at

$$\Delta V_{x,k} \le x_k^T \left(2A_c^T \Lambda_1 A_c - \Lambda_1\right) x_k + 4\left[B\tilde{W}_k^T \Phi(\bar{x}_k)\bar{u}_k\right]^T \Lambda_1 \left[B\tilde{W}_k^T \Phi(\bar{x}_k)\bar{u}_k\right] + 4\left[B\Xi_k \bar{u}_k\right]^T \Lambda_1 \left[B\Xi_k \bar{u}_k\right].$$

$$\leq -\sigma_{\min}(Q)\|x_k\|^2 + 4\Phi_{\max}^2\|\Lambda_1\|\|\tilde{W}_k\|^2\|\bar{u}_k\|^2 + 4\|\Lambda_1\|\|\bar{u}_k\|^2\Xi_{\max}^2.$$

where $2A_c^T\Lambda_1 A_c - \Lambda_1 = -Q$ and $\|B\| = 1$.

Remark A.1: The Lyapunov equation $\bar{A}_c^T\Lambda_1\bar{A}_c - \Lambda_1 = -Q$ has a positive definite solution only when the matrix $\bar{A}_c = \sqrt{2}A_c$ is Schur. As per the definition of matrix A_c in (5), the control gain K can be selected to ensure $\bar{A}_c$ is Schur.

By using the facts $\boldsymbol{\varphi}_{f,\max} \leq \Phi_{\max}$ and $\|W_{f,k}\| \leq \|W_k\|$, the control input at the trigger instants given in (20) satisfies

$$\begin{aligned}\|\bar{u}_k\|^2 &= 1 + \|u_k\|^2 = 1 + \left\|[-\hat{W}_{f,k}^T\varphi_f(\bar{x}_k) + Kx_k]/\hat{W}_{g,k}^T\varphi_g(\bar{x}_k)\right\|^2 \\ &\leq 1 + \frac{4\Phi_{\max}^2 W_{\max}^2}{g_{\min}^2} + \frac{4\Phi_{\max}^2}{g_{\min}^2}\|\tilde{W}_k\|^2 + \frac{2\|K\|^2}{g_{\min}^2}\|x_k\|^2.\end{aligned} \tag{A.9}$$

Substituting the inequality (A.9) and separating the cross product term using Young's inequality, $2ab \leq a^2 + b^2$, the first difference is bound by

$$\begin{aligned}\Delta V_{x,k} \leq\ & -\sigma_{\min}(Q)\|x_k\|^2 + 4\Phi_{\max}^2\|\Lambda_1\|\|\tilde{W}_k\|^2 + (16/g_{\min}^2)\Phi_{\max}^4 W_{\max}^2\|\Lambda_1\|\|\tilde{W}_k\|^2 \\ &+ (20/g_{\min}^2)\Phi_{\max}^4\|\Lambda_1\|\|\tilde{W}_k\|^4 + (4/g_{\min}^2)\|\Lambda_1\|\|K\|^4\|x_k\|^4 + (16/g_{\min}^2)\|\Lambda_1\|\Phi_{\max}^2\Xi_{\max}^2 W_{\max}^2 \\ &+ (16/g_{\min}^2)\|\Lambda_1\|\Phi_{\max}^2\Xi_{\max}^2\|\tilde{W}_k\|^2 + (8/g_{\min}^2)\|\Lambda_1\|\Xi_{\max}^2\|K\|^2\|x_k\|^2 + 4\|\Lambda_1\|\Xi_{\max}^2.\end{aligned} \tag{A.10}$$

Considering the second term of the Lyapunov function $V_{\hat{x},k} = \hat{x}_k^T\Lambda_2\hat{x}_k$,, the first difference along the closed-loop SE dynamics (32) with $\hat{x}_k = x_k$ at $k = k_i$ becomes

$$\begin{aligned}\Delta V_{\hat{x},k} &= \hat{x}_{k+1}^T\Lambda_2\hat{x}_{k+1} - \hat{x}_k^T\Lambda_2\hat{x}_k = (A_c x_k)^T\Lambda_2(A_c x_k) - x_k^T\Lambda_2 x_k \\ &= x_k^T(A_c^T\Lambda_2 A_c - \Lambda_2)x_k = -x_k^T\bar{Q}x_k \leq -\sigma_{\min}(\bar{Q})\|x_k\|^2.\end{aligned} \tag{A.11}$$

Moving on for the third term, $V_{A,k} = (\hat{x}_k^T \Lambda_2 \hat{x}_k)^2$, the first difference $\Delta V_{A,k} = (\hat{x}_{k+1}^T \Lambda_2 \hat{x}_{k+1})^2 - (\hat{x}_k^T \Lambda_2 \hat{x}_k)^2$ can be written as $\Delta V_{A,k} = \Delta V_{\hat{x},k} \left(\Delta V_{\hat{x},k} + 2 x_k^T \Lambda_2 x_k \right)$. Substituting $\Delta V_{\hat{x},k}$ from (A.11) reveals that

$$\Delta V_{A,k} \le -\sigma_{\min}(\bar{Q}) \sigma_{\min}(A_c^T \Lambda_2 A_c + \Lambda_2) \left\| x_k \right\|^4 . \tag{A.12}$$

Now, the first difference of the fourth term $V_{\tilde{W},k}$ in the Lyapunov function can be written from (A.2) in Lemma 1 and given by

$$\Delta V_{\tilde{W},k} \le -\beta \left\| \tilde{W}_k \right\|^2 + B_M^{\tilde{W}} . \tag{A.13}$$

Considering the last term $V_{B,k} = tr\{\tilde{W}_k^T \tilde{W}_k\}^2$, the first difference can be computed using (A.13) as follows

$$\Delta V_{B,k} = \left(tr\{\tilde{W}_{k+1}^T \tilde{W}_{k+1}\}^2 - tr\{\tilde{W}_k^T \tilde{W}_k\}^2 \right) \le (-\beta \left\| \tilde{W}_k \right\|^2 + B_M^{\tilde{W}})((2-\beta) \left\| \tilde{W}_k \right\|^2 + B_M^{\tilde{W}}).$$

Appling Young's inequality $2ab \le pa^2 + \left(b^2 / p \right)$ reveals that

$$\Delta V_{B,k} \le -(1/2)\beta(2-\beta) \left\| \tilde{W}_k \right\|^4 + \left(((2-\beta)/2\beta) + 1 \right) (B_M^{\tilde{W}})^2, \tag{A.14}$$

where $(2-\beta) > 0$ by the selection of $0 < \kappa < 1/4$.

Finally, combining all the individual first differences given in (A.10), (A.11), (A.12), (A.13), and (A.14) the overall first difference $\Delta V_k = \Delta V_{x,k} + \eta \Delta V_{\hat{x},k} + \varpi \Delta V_{A,k} + \Psi \Delta V_{\tilde{W},k} + \Pi \Delta V_{B,k}$, with η, ϖ, Ψ and Π from (A.8), found to be

$$\begin{aligned} \Delta V_k \le & -\sigma_{\min}(Q) \left\| x_k \right\|^2 - \left(1/g_{\min}^2 \right) \left\| \Lambda_1 \right\| \Xi_{\max}^2 \left\| K \right\|^2 \left\| x_k \right\|^2 \\ & - \left(4\Phi_{\max}^2 \left\| \Lambda_1 \right\| + \frac{16}{g_{\min}^2} \left\| \Lambda_1 \right\| \Phi_{\max}^2 \left(\Phi_{\max}^2 W_{\max}^2 + \Xi_{\max}^2 \right) \right) \left\| \tilde{W}_k \right\|^2 \\ & - \left(1/g_{\min}^2 \right) \left\| \Lambda_1 \right\| \left\| K \right\|^4 \left\| x_k \right\|^4 - \left(1/g_{\min}^2 \right) \Phi_{\max}^4 \left\| \Lambda_1 \right\| \left\| \tilde{W}_k \right\|^4 + B_{TM}^{c2} , \end{aligned} \tag{A.15}$$

where $B_{TM}^{c2} = \left(16/g_{\min}^2\right)\|\Lambda_1\|\Phi_{\max}^2\Xi_{\max}^2 W_{\max}^2 + 4\|\Lambda_1\|\Xi_{\max}^2 + \Psi B_M^{\tilde{W}} + \Pi\left(((2-\beta)/2\beta)+1\right)(B_M^{\tilde{W}})^2$.

From (A.15), the Lyapunov first difference ΔV_k is less than zero as long as

$$\|x_k\| > \max\left\{\sqrt{B_{TM}^{c2}/\sigma_{\min}(Q)}, \sqrt{g_{\min}^2 B_{TM}^{c2}/\|\Lambda_1\|\Xi_{\max}^2\|K\|^2}, \sqrt[4]{g_{\min}^2 B_{TM}^{c2}/\|\Lambda_1\|\|K\|^4}\right\} \equiv B_{1,M}^{x} \text{ or}$$

$$\|\tilde{W}_k\| > \max\left\{\sqrt{g_{\min}^2 B_{TM}^{c2}/\{4\Phi_{\max}^2 g_{\min}^2\|\Lambda_1\| + 16\|\Lambda_1\|\Phi_{\max}^2(\Phi_{\max}^2 W_{\max}^2 + \Xi_{\max}^2)\}}, \sqrt[4]{g_{\min}^2 B_{TM}^{c2}/\Phi_{\max}^4\|\Lambda_1\|}\right\} \equiv B_{1,M}^{\tilde{W}}.$$

Therefore, by using the Lyapunov theorem [20], the system state x_k, SE state $\hat{x}_k$, and the NN weight estimation error $\tilde{W}_k$ are bounded at the trigger instants. Further, when $i > p$ or, alternatively, all trigger instants $k_i > k_p$ where p is an integer representing the events, the system state x_k, SE state $\hat{x}_k$, and the NN weight estimation error $\tilde{W}_k$ are all ultimately bounded.

Case II: During the inter-event times ($k_i < k < k_{i+1}$, $\forall i \in \mathbb{N}$)

Consider the Lyapunov function given in (A.8) in Case I. Similar to Case I, we will evaluate the individual terms separately. Note that the NN weights are not updated during the inter-event times and held at their previous values.

Consider the first term $V_{x,k} = x_k^T \Lambda_1 x_k$ of the Lyapunov function candidate (A.8). The first difference $\Delta V_{x,k}$ along the closed-loop system trajectory (30) can be expressed as

$$\begin{aligned}\Delta V_{x,k} &= x_{k+1}^T \Lambda_1 x_{k+1} - x_k^T \Lambda_1 x_k \\ &= [A_c x_k + BKe_k^s + B\tilde{W}_k^T \Phi(\bar{x}_k)\bar{u}_k + B\Xi_k \bar{u}_k + B\hat{W}_k^T \tilde{\Theta}\left(\bar{x}_k, \hat{\bar{x}}_k\right)\bar{u}_k]^T \Lambda_1 [A_c x_k + BKe_k^s \\ &+ B\tilde{W}_k^T \Phi(\bar{x}_k)\bar{u}_k + B\Xi_k \bar{u}_k + B\hat{W}_k^T \tilde{\Theta}\left(\bar{x}_k, \hat{\bar{x}}_k\right)\bar{u}_k] - x_k^T \Lambda_1 x_k.\end{aligned}$$

Applying C-S inequality, the first difference can be represented as

$$\Delta V_{x,k} \le -x_k^T Q x_k + 8\left(B\tilde{W}_k^T \Phi(\bar{x}_k)\bar{u}_k\right)^T \Lambda_1 \left(B\tilde{W}_k^T \Phi(\bar{x}_k)\bar{u}_k\right) + 8\left(B\Xi_k \bar{u}_k\right)^T \Lambda_1 \left(B\Xi_k \bar{u}_k\right)$$
$$+4\left[BKe_k^s + B\hat{W}_k^T \tilde{\Theta}\left(\bar{x}_k, \hat{\bar{x}}_k\right)\bar{u}_k\right]^T \Lambda_1 \left[BKe_k^s + B\hat{W}_k^T \tilde{\Theta}\left(\bar{x}_k, \hat{\bar{x}}_k\right)\bar{u}_k\right],$$

where Q satisfies the Lyapunov equation $\bar{A}_c^T \Lambda_1 \bar{A}_c - \Lambda_1 = -Q$ with $\bar{A}_c = \sqrt{2}A_c$. By using Frobenius norm and triangle inequality with the fact $\|B\| = 1$ reveals

$$\Delta V_{x,k} \le -\sigma_{\min}(Q)\|x_k\|^2 + 8\Phi_{\max}^2 \left\|\tilde{W}_k\right\|^2 \|\bar{u}_k\|^2 \|\Lambda_1\| + 8\|\bar{u}_k\|^2 \|\Lambda_1\| \|\Xi_k\|^2$$
$$+4\left\|BKe_k^s + B\hat{W}_k^T \tilde{\Theta}\left(\bar{x}_k, \hat{\bar{x}}_k\right)\bar{u}_k\right\|^2 \|\Lambda_1\|.$$

Applying C-S and Young's inequalities and replacing $\left\|\tilde{\Theta}\left(\bar{x}_k, \hat{\bar{x}}_k\right)\right\| \le L\left\|\bar{x}_k - \hat{\bar{x}}_k\right\| \equiv L_\Phi \left\|e_k^s\right\|$ from Assumption 3, the first difference can be expressed as

$$\Delta V_{x,k} \le -\sigma_{\min}(Q)\|x_k\|^2 + 8\Phi_{\max}^2 \left\|\tilde{W}_k\right\|^2 \|\bar{u}_k\|^2 \|\Lambda_1\| + 8\|K\|^2 \left\|e_k^s\right\|^2 \|\Lambda_1\|$$
$$+8L_\Phi^2 \left\|\hat{W}_k\right\|^2 \left\|e_k^s\right\|^2 \|\bar{u}_k\|^2 \|\Lambda_1\| + 8\|\bar{u}_k\|^2 \|\Lambda_1\| \|\Xi_k\|^2 \quad \text{(A.16)}$$

By definition of the control input (20) for $k_i < k < k_{i+1}$, the following inequality holds.

$$\|\bar{u}_k\|^2 \le 1 + \frac{4\Phi_{\max}^2 W_{\max}^2}{g_{\min}^2} + \frac{4\Phi_{\max}^2}{g_{\min}^2}\left\|\tilde{W}_k\right\|^2 + \frac{2\|K\|^2}{g_{\min}^2}\|\hat{x}_k\|^2. \quad \text{(A.17)}$$

Substituting (A.17) in the first difference (A.16) and with simple mathematical manipulation one can reach at

$$\Delta V_{x,k} \le -\sigma_{\min}(Q)\|x_k\|^2 + \left(16/g_{\min}^2\right)\|K\|^2 \Xi_{\max}^2 \|\Lambda_1\| \|\hat{x}_k\|^2 + \left(8/g_{\min}^2\right)\|K\|^2 \|\Lambda_1\| \|\hat{x}_k\|^4 \Phi_{\max}^2$$
$$+\left(8\|K\|^2 \|\Lambda_1\| + 8L_\Phi^2 \left\|\hat{W}_k\right\|^2 \|\bar{u}_k\|^2 \|\Lambda_1\|\right)\left\|e_k^s\right\|^2 + 8\Phi_{\max}^2 \|\Lambda_1\|\left(1 + \left(4/g_{\min}^2\right)\left(\Phi_{\max}^2 W_{\max}^2 + \Xi_{\max}^2\right)\right)\left\|\tilde{W}_k\right\|^2$$
$$+\left(8/g_{\min}^2\right)\Phi_{\max}^2 \|\Lambda_1\| \left\|\tilde{W}_k\right\|^4 \left(4\Phi_{\max}^2 + \|K\|^2\right) + 32\left(\Phi_{\max}^2/g_{\min}^2\right)\|\Lambda_1\| W_{\max}^2 \Xi_{\max}^2 + 8\|\Lambda_1\| \Xi_{\max}^2.$$

Recall the event-triggering condition (33) for the case when the system state vector is outside the ultimate bound. During the inter-event times the inequality (33) holds. Substituting in the above first difference one can arrive at

$$\begin{aligned}\Delta V_{x,k} \le & -(1-\Gamma)\sigma_{\min}(Q)\|x_k\|^2 + (16/g_{\min}^2)\|K\|^2 \Xi_{\max}^2 \|\Lambda_1\| \times \|\hat{x}_k\|^2 \\ & +8\Phi_{\max}^2 \|\Lambda_1\| \|\tilde{W}_k\|^2 \left(1+(4/g_{\min}^2)(\Phi_{\max}^2 W_{\max}^2 + \Xi_{\max}^2)\right) \\ & +(8/g_{\min}^2)\|K\|^2 \|\Lambda_1\| \Phi_{\max}^2 \|\hat{x}_k\|^4 + (8/g_{\min}^4)\Phi_{\max}^2 \|\Lambda_1\| \|\tilde{W}_k\|^4 \left(4\Phi_{\max}^2 + \|K\|^2\right) \\ & +\left(8\|\Lambda_1\| + (32/g_{\min}^4)\Phi_{\max}^2 W_{\max}^2 \|\Lambda_1\|\right)\Xi_{\max}^2 .\end{aligned} \tag{A.18}$$

Consider the second term $V_{\hat{x},k} = \hat{x}_k^T \Lambda_2 \hat{x}_k$ of the Lyapunov function (A.8). The first difference $\Delta V_{\hat{x},k}$ along the closed-loop SE dynamics (32) for $k_i < k < k_{i+1}$ can be represented as

$$\Delta V_{\hat{x},k} = \hat{x}_{k+1}^T \Lambda_2 \hat{x}_{k+1} - \hat{x}_k^T \Lambda_2 \hat{x}_k = -\hat{x}_k^T \bar{Q} \hat{x}_k \le -\sigma_{\min}(\bar{Q})\|\hat{x}_k\|^2, \tag{A.19}$$

where the positive definite matrix $\bar{Q}$ satisfies the Lyapunov equation $A_c^T \Lambda_2 A_c - \Lambda_2 = -\bar{Q}$.

The first difference of the third term $V_{A,k} = \left(\hat{x}_k^T \Lambda_2 \hat{x}_k\right)^2$ can be written using (A.19) as

$$\Delta V_{A,k} \le -\hat{x}_k^T \bar{Q} \hat{x}_k \left(-\hat{x}_k^T \bar{Q} \hat{x}_k + 2\hat{x}_k^T \Lambda_2 \hat{x}_k\right) \le -\sigma_{\min}(\bar{Q})\sigma_{\min}\left(A_c^T \Lambda_2 A_c + \Lambda_2\right)\|\hat{x}_k\|^4 . \tag{A.20}$$

The first difference of the fourth term, $V_{\tilde{W},k} = tr\{\tilde{W}_k^T \tilde{W}_k\}$, in (A.8) can be written from (A.3) and given as

$$\Delta V_{\tilde{W},k} = 0 \tag{A.21}$$

Therefore, the first difference of $V_{B,k} = tr\{\tilde{W}_k^T \tilde{W}_k\}^2$ from (A.21) is written as

$$\Delta V_{B,k} = 0 . \tag{A.22}$$

The overall first difference, $\Delta V_k = \Delta V_{x,k} + \eta \Delta V_{\hat{x},k} + \varpi \Delta V_{A,k} + \Psi \Delta V_{\tilde{W},k} + \Pi \Delta V_{B,k}$, by combining (A.18), (A.19) , (A.20), (A.21), and (A.22), and recalling the definition of η and ϖ, is upper bounded by

$$\begin{aligned} \Delta V_k \le & -\sigma_{\min}(Q)(1-\Gamma)\|x_k\|^2 - (1/g_{\min}^2)\Phi_{\max}^2 \|K\|^2 \|\Lambda_1\| \|\hat{x}_k\|^4 \\ & -(1/g_{\min}^2 \|K\|^2 \Xi_{\max}^2 \|\Lambda_1\| \|\hat{x}_k\|^2 + B_{\Xi}^M + B_{\tilde{W},i}^M, \end{aligned} \tag{A.23}$$

where $B_{\tilde{W},i}^M = 8\Phi_{\max}^2 \|\Lambda_1\| (1 + (4/g_{\min}^2)(\Phi_{\max}^2 \mathbf{W}_{\max}^2 + \Xi_{\max}^2)) B_i^{\tilde{W}} + (8\Phi_{\max}^2 \|\Lambda_1\| / g_{\min}^2)(4\Phi_{\max}^2 + \|K\|^2) B_i^{\tilde{W}^2}$ and $B_{\Xi}^M = \left(8\|\Lambda_1\| + (32/g_{\min}^4)\Phi_{\max}^2 W_{\max}^2 \|\Lambda_1\|\right)\Xi_{\max}^2$ with $B_i^{\tilde{W}}$ is piecewise constant bound of $\tilde{W}_k$ for i^{th} inter-event time from Lemma1. From (A.23) the overall first difference ΔV_k is less than zero as long as

$$\|x_k\| > \sqrt{\left(B_{\Xi}^M + B_{\tilde{W},i}^M\right) / \sigma_{\min}(Q)(1-\Gamma)} \equiv B_{2,i}^x \text{ or}$$

$$\|\hat{x}_k\| > \max\left\{ \sqrt{\left(B_{\Xi}^M + B_{\tilde{W},i}^M\right) g_{\min}^2 / \|K\|^2 \Xi_{\max}^2 \|\Lambda_1\|}, \sqrt[4]{\left(B_{\Xi}^M + B_{\tilde{W},i}^M\right) g_{\min}^2 / \|K\|^2 \|\Lambda_1\| \Phi_{\max}^2} \right\} \equiv B_{2,i}^{\hat{x}}.$$

This implies, either the system state vector outside the ball of radius $B_{2,i}^x$ or the SE state vector outside the ball of radius $B_{2,i}^{\hat{x}}$, both will converge to their respective bounds in a finite time. Since, inter-event times are followed by the trigger instants, the initial values of x_k, $\hat{x}_k$, and $\tilde{W}_k$, during the inter-event times are the updated values from the trigger instants. It is shown in Case I that x_k, $\hat{x}_k$, and $\tilde{W}_k$, are bounded at the trigger instants. Therefore, the system and SE state vectors are bounded during the inter-event times. Note that the function $B_{\tilde{W},i}^M$ in (A.23) is a piecewise constant function since $B_i^{\tilde{W}}$ in (A.7), from Lemma 1, is constant during i^{th} inter-event time. Therefore, the

bounds for the system and the SE state vectors, $B_{2,i}^{x}$, and $B_{2,i}^{\hat{x}}$, respectively, are piece wise constant functions.

During initial learning phase of the NN, the upper bound on the NN weight estimation error $B_i^{\tilde{W}}$ in (A.7) may be large. Hence, the piecewise constant function $B_{\tilde{W},i}^{M}$ and in turn $B_{2,i}^{x}$, and $B_{2,i}^{\hat{x}}$ are of larger value. The system and SE state vectors inside the ball of radius $B_{2,i}^{x}$, and $B_{2,i}^{\hat{x}}$, respectively, may increase within these bounds. It follows that the Lyapunov function (A.8) may increase and bounded by the piecewise constant bound. The upper bound on the Lyapunov function using the upper bounds of the system state, SE state and NN weight estimation error can be expressed as

$$V_k \le B_{1,i}^{x^2} + \eta B_{2,i}^{\hat{x}^2} + \varpi B_{2,i}^{\hat{x}^4} + \Psi B_i^{\tilde{W}} + \Pi B_i^{\tilde{W}^2}, \tag{A.24}$$

for $k_i < k < k_{i+1}$, $\forall i \in \mathbb{N}$.

To show the UB of x_k, $\hat{x}_k$, and $\tilde{W}_k$, we need to show the functions $B_{\tilde{W},i}^{M}$, $B_{2,i}^{x}$, and $B_{2,i}^{\hat{x}}$ converge to their ultimate values. The bounds $B_{\tilde{W},i}^{M}$, $B_{2,i}^{x}$, and $B_{2,i}^{\hat{x}}$ are functions of $B_i^{\tilde{W}}$. Since, $B_i^{\tilde{W}}$ in (A.7) is a converging sequence, shown in Lemma 1, and converges to $B_{UB}^{\tilde{W}}$ for all $i > p$, the function $B_{\tilde{W},i}^{M}$ in (A.23) converges to the ultimate value, i.e., $B_{\tilde{W},i}^{M} \to B_{\tilde{W},M}^{M}$ for all $i > p$ where $B_{\tilde{W},M}^{M} = 8\Phi_{\max}^2 \|\Lambda_1\|(1+(4/g_{\min}^2)(\Phi_{\max}^2 W_{\max}^2 + \Xi_{\max}^2))$ $\times B_{UB}^{\tilde{W}} + (8\Phi_{\max}^2 \|\Lambda_1\|/g_{\min}^2)(4\Phi_{\max}^2 + \|K\|^2) B_{UB}^{\tilde{W}^2}$. Consequently, the bounds $B_{2,i}^{x} \to B_{2,M}^{x}$, and $B_{2,i}^{\hat{x}} \to B_{2,M}^{\hat{x}}$ for all $i > p$, where $B_{2,M}^{x} = \sqrt{\left(B_{\Xi}^{M} + B_{\tilde{W},M}^{M}\right)/\sigma_{\min}(Q)(1-\Gamma)}$ and $B_{2,M}^{\hat{x}} = \max\left\{\sqrt{\left(B_{\Xi}^{M} + B_{\tilde{W},M}^{M}\right) g_{\min}^2 / \|K\|^2 \Xi_{\max}^2 \|\Lambda_1\|}, \sqrt[4]{\left(B_{\Xi}^{M} + B_{\tilde{W},M}^{M}\right) g_{\min}^2 / \|K\|^2 \|\Lambda_1\| \Phi_{\max}^2}\right\}$.

Therefore, combining results from Case I and Case II, the system state x_k, SE state $\hat{x}_k$, and the NN weight estimation error $\tilde{W}_k$ are bounded for all time and converge to the ultimate bound when $i > p$ or with events occurring such that $k_i > k_p$. Therefore, all the closed-loop system signals are UB for all time $k > k_0 + N$ since k_i is a subsequence of k where $N \geq k_p$ represents the time instant.

From both the cases of the proof and Lemma 1, the bounds for the system state vector, SE state vector and NN weight estimation error can be selected as $B_x = \max(B_{1,M}^x, B_{2,M}^x)$, $B_{\hat{x}} = \max(B_{1,M}^x, B_{2,M}^{\hat{x}})$, and $B_{\tilde{W}} = \max(B_{UB}^{\tilde{W}}, B_{2,M}^{\tilde{W}})$, respectively.

Remark A.2: It is routine to check that for the case $\left|\hat{g}\left(\hat{x}_k\right)\right| < g_{\min}$, in (21), the first differences in (A.15) and (A.23) also hold. Therefore, with similar arguments the closed-loop event-triggered system is ultimately bounded. ■

III. NEAR OPTIMAL EVENT-TRIGGERED CONTROL OF NONLINEAR DISCRETE-TIME SYSTEMS USING NEURO DYNAMICS PROGRAMMING

Avimanyu Sahoo, Hao Xu and S. Jagannathan

Abstract *— This paper presents an event-triggered near optimal control of uncertain nonlinear discrete time systems. Event driven neuro-dynamic programming is utilized to design the control policy. A neural network (NN) based identifier, with event-based state and input vectors, is utilized to learn the system dynamics. An actor-critic framework is used to learn the cost function and the optimal control input. The NN weights of the identifier, critic and actor NNs are tuned once every triggered instant with aperiodic update laws. An adaptive event-triggering condition to decide the trigger instants is derived. Thus, a suitable number of events are generated to ensure a desired accuracy of approximation. A near optimal performance is achieved without value and/or policy iterations. A detailed analysis of nontrivial inter-event time is presented. An explicit formula to show the reduction in computation is also derived. The Lyapunov technique is used in conjunction with the event-triggering condition to guarantee ultimate boundedness of the closed-loop system. Simulation results are included to verify the performance of the controller. The net result is the development of event-driven neuro dynamic programming.*

Index Terms — Event-triggered control, Hamilton-Jacobi-Bellman equation, neuro-dynamic programming, neural networks, optimal control.

1. INTRODUCTION

Event-triggered control (ETC) [1]-[7] is evolved as an alternate control paradigm in the recent times. This control paradigm is found to be effective in terms of resource utilization. The ETC scheme uses events to sample the system state and execute the controller in an aperiodic manner. The aperiodic sampling and execution reduces computational costs for the closed-loop system. In the case of a networked control system (NCS) [8], the ETC scheme saves network bandwidth due to the event-based aperiodic transmissions. The sampling and transmission instants, referred to as event-trigger instants, are decided by using a state dependent criterion. The threshold in the criterion is designed analytically via the Lyapunov stability technique. Thus, the event-triggered paradigm saves resources, and maintains both stability and closed-loop performance.

Recently, various event-triggered control schemes [1]-[5] have been introduced in the literature for linear [3], [5] and nonlinear systems [2], [4]. Typically, in the event-triggered control schemes [2]-[5], system dynamics are considered either completely known [2], [4]-[5], or with a small uncertainty [3]. In contrast, in our previous work [6], [7] an attempt has been made to design an event-based controller for systems with uncertain dynamics. In [6], the knowledge of system dynamics is partially relaxed by using an event-based neural network (NN) approximator. The NN based design is extended to the case of completely unknown dynamics in [7]. In both cases, the state dependent criteria, referred to as event-trigger conditions, are made adaptive. This is in contrast with traditional non-adaptive event-trigger conditions [2]-[4] in the literature. These adaptive criteria generated a required number of events during initial online

learning phase of NN. This facilitated the event-based approximation of the unknown dynamics with aperiodic weight update. A trade-off is observed between the accuracy of NN approximation and reduction in computation. However, these designs [6]-[7] have only considered stability without having any performance index to optimize.

The traditional optimal control design approach [9] has also been studied in an event-triggered control context by the authors in [10]-[12]. The authors in [10] studied the optimal ETC in a constrained communication scenario by using the certainty equivalence principle. Further, the authors in [11] extended the results to an event-triggered context with the help of a linear quadratic Gaussian (LQG) approach. The separation principle is used to design the optimal control input and the optimal transmission instants. However, these methods [10]-[11] use backward-in-time Riccati equation (RE) based solution with completely known system dynamics.

Traditionally, adaptive dynamic programming (ADP) or neuro-dynamic programming (NDP) [13] techniques are used to design the optimal control policy in a forward-in-time and online manner. These techniques use the policy and/or value iterations to solve the Hamilton-Jacobi-Bellman (HJB) equation online. However, a significant number of iterations within a sampling interval are needed to maintain system stability resulting in high computational cost. Further, the knowledge of the control coefficient function is also necessary to compute the optimal control policy.

For a finite-time [14] optimal control, the solution to the HJB equation (i.e., the cost function) becomes explicitly time varying. The terminal cost constraint must also be satisfied at the same time. The event-based sampling of the state vector and uncertain

system dynamics complicate the problem further. Therefore, NDP over the finite-horizon becomes more involved than in the infinite horizon case.

Motivated by the above limitations, in this paper, we propose a novel NDP technique to solve the fixed final-time optimal control. An event-triggered uncertain nonlinear discrete-time system is considered for the purpose of design. The proposed approach functions in a forward-in-time and online manner. Two NNs, in an actor-critic NN [15] framework, are used to approximate the time-varying cost function and the optimal control input. A NN identifier is used to relax the complete knowledge of system dynamics. A novel adaptive event-trigger condition is also developed which not only reduces the number of controller updates but also facilitates the NN approximation.

Aperiodic NN tuning laws are introduced to update the identifier, actor and critic NN weights. The NN weights are updated once a triggered instant and held during the inter-event duration. These aperiodic updates reduce the computation when compared to a traditional NN based schemes [16]. The Lyapunov direct method as in [4], [17] is used to prove the ultimate boundedness (UB) of the closed-loop event-triggered system.

The contributions of this paper include: 1) the design of event-triggered finite-time optimal control scheme for an uncertain nonlinear discrete time system, 2) the design of a novel adaptive event-trigger condition, 3) the development of aperiodic tuning laws to save computation, and 4) the demonstration of the closed-loop stability by using the Lyapunov technique.

The rest of the paper is organized as follows. Section 2 presents the background along with the problem statement. Section 3 details the finite horizon event-based optimal control design. The main results are claimed in Section 4 and non-triviality of the inter-

event times is discussed in Section 5. Simulation results are included in Section 6. Conclusions are drawn in Section 7. The appendix contains the detailed proof of the lemmas and the theorems.

2. BACKGROUND AND PROBLEM FORMULATION

In this section, we present a brief background on the ETC. Subsequently, the near optimal control design is formulated. A discussion on the extension of NN approximation to event-based sampling is also presented.

2.1 BACKGROUND ON ETC

Consider the uncertain nonlinear discrete-time system represented as

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

where $x_k \in \Re^n$ and $u_k \in \Re^m$ represent the system state and the control input vectors, respectively. The smooth functions $f(x_k) \in \Re^n$ and $g(x_k) \in \Re^{n\times m}$ denote the system dynamics that are considered unknown. Let the equilibrium point $x = 0$ be unique in a compact set for all $x_k \in \mathcal{D}_x \subset \Re^n$. The following standard assumption is necessary in order to proceed.

Assumption 1: The system (1) is controllable and observable. The unknown control coefficient matrix $g(x_k)$ is bounded for all $x_k \in \mathcal{D}_x \subset \Re^n$ such that $\|g(x_k)\| \le g_M$ where $g_M > 0$ is a known, positive constant. The state vector is available for measurement.

In the event-triggered formalism, the system state vector x_k is released and the controller is updated only when an event occurs. Hence, zero-order-holds (ZOH) are used to retain the last event-sampled state and the control input vectors until the next arrives. The error between the current measured state vector, x_k, and the state vector at the ZOH, $\breve{x}_k$, is referred to as event-trigger error. It is defined by

$$e_{ET,k} = x_k - \breve{x}_k. \tag{2}$$

The event-trigger error (2) is used to determine the event-trigger instants by comparing it with a state dependent threshold. A monotonically increasing subsequence of time instants $\{k_i\}_{i=1}^{\infty}$ with $k_0 = 0$ can be defined as the event-trigger instants. The last held state vector, $\breve{x}_k$, at ZOH is updated at each $k = k_i$ for $i = 1, 2, \cdots$ with the current system state. Thus, the last held state vector can be written as

$$\breve{x}_k = x_{k_i}, \ k_i \le k < k_{i+1}, i = 1, 2, \cdots. \tag{3}$$

In an event-based framework, the control input can be described as

$$u_k = \upsilon(\breve{x}_k), \ \ k_i \le k < k_{i+1}, \ \forall \ \ i = 1, 2, \cdots, \tag{4}$$

where $\upsilon(\breve{x}_k)$ is a function of the event-based state vector. Next, the problem for the finite horizon optimal control in an event-based scenario is formulated.

2.2 PROBLEM FORMULATION

Our primary objective is to design a sequence of control inputs, u_k to minimize a time-varying cost function in an ETC framework. The cost function is given by

$$V(x_k, k) = \psi(x_N, N) + \sum_{j=k}^{N-1} r(x_j, u_j, j), \tag{5}$$

where $r(x_j, u_j, j) = Q(x_j, j) + u_j^T R u_j$ is the cost-to-go in the interval of interest $j \in [k, N]$. The function $Q(x_k, k) \in \Re$ is a positive definite function that penalizes the system state, x_k. The matrix $R \in \Re^{m \times m}$ is a positive definite matrix that penalizes the control input, u_k. The terminal cost $\psi(x_N, N)$ penalizes the terminal state x_N where N is the terminal time instant. For finite horizon case, the cost-to-go, $r(x_k, u_k, k)$, depends

explicitly on time k in the interval of interest $[k, N]$. Therefore, the control input also becomes time varying.

Assumption 2: The initial control input, u_0, is admissible [15] to keep the cost function finite.

The terminal cost for the finite horizon cost function (5) can be written as

$$V(x_N, N) = \psi(x_N, N), \tag{6}$$

where $V(x_N, N)$ is the cost at the terminal time N. The cost function can also be rewritten as

$$V(x_k, k) = r(x_k, u_k, k) + \sum_{j=k+1}^{N-1} \{r(x_j, u_j, j)\} + V\left(x_N, N\right)$$

$$= r(x_k, u_k, k) + V(x_{k+1}, k+1), \tag{7}$$

where $V(x_{k+1}, k+1) = V\left(x_N, N\right) + \sum_{j=k+1}^{N-1} r(x_j, u_j, j)$ is the cost function from the time instant $k+1$ onwards. According to Bellman's principle of optimality, the optimal cost, $V^*(x_k, k)$ satisfies the discrete-time Hamilton-Jacobi-Bellman (HJB) equation. It is given by

$$V^*(x_k, k) = \min_{u_k} \left\{ r(x_k, u_k, k) + V^*(x_{k+1}, k+1) \right\}, \tag{8}$$

where $V^*(x_k, k)$ is the optimal cost at the time instant k and $V^*(x_{k+1}, k+1)$ is the optimal cost for $k+1$ onwards. The optimal control sequence u_k^* can be derived by using stationarity condition [9] and written as

$$u_k^* = -(1/2) R^{-1} g^T\left(x_k\right) \partial V^*\left(x_{k+1}, k+1\right) / \partial x_{k+1}. \tag{9}$$

The optimal control policy (9) depends explicitly on the solution of the HJB equation, i.e., the optimal cost $V^*(x_k,k)$. The control policy is also a function of control coefficient function $g(x_k)$ and the state vector x_{k+1} at the time instant k.

It is practically almost impossible to find an analytical solution of the HJB equation. Therefore, approximation based techniques (NDP) are used to solve the HJB equation. In this paper, actor and critic NNs are utilized to approximate both the optimal control policy and cost function, respectively. The approximations are carried out with the event-based availability of the system state vector. Hence, the universal approximation property of the NNs is revisited with an extension to event-based approximation.

2.3 NN APPROXIMATION WITH EVENT BASED SAMPLING

The universal approximation property [16] of NN can be extended to achieve a desired level of accuracy with event-based availability of the state vector in (3). The theorem introduced next extends the approximation property of NNs for event-based sampling.

Theorem 1: Let $h(x_k,k)\in\Re^n$ be a smooth and continuous function in a compact set for all $x\in\mathcal{D}_x\subset\Re^n$. Then, there exists a NN with a sufficient number of neurons such that $h(x_k,k)$ can be approximated with event sampled inputs. Further, the function $h(x_k,k)$ with constant weights and event-based time-varying activation function is given by

$$h(x_k,k)=W^T\sigma(\breve{x}_k,k)+\varepsilon_e(\breve{x}_k,e_{ET,k},k), \tag{10}$$

where $W\in\Re^{l\times n}$ is the constant unknown target weight matrix. The number of hidden-layer neurons denoted by l while $\sigma(\breve{x}_k,k)\in\Re^l$ is a bounded event-based time varying

activation function. The event-based NN reconstruction error is denoted by $\varepsilon_e(\breve{x}_k, e_{ET,k}, k) = W^T[\sigma(\vartheta(\breve{x}_k, e_{ET,k}), k) - \sigma(\breve{x}_k, k)] + \varepsilon(\vartheta(\breve{x}_k, e_{ET,k}), k)$ where $\vartheta(\breve{x}_k, e_{ET,k}) = \breve{x}_k + e_{ET,k}$. The function $\sigma(\vartheta(\breve{x}_k, e_{ET,k}), k) = \sigma(x_k, k)$ is the periodic time-based activation function, $\varepsilon(\vartheta(\breve{x}_k, e_{ET,k}), k) = \varepsilon(x_k, k)$ is the traditional reconstruction error, and $\breve{x}_k$ is the latest available event-sampled state.

Proof: Refer to the Appendix.

Remark 1: The event-based reconstruction error $\varepsilon_e(\breve{x}_k, e_{ET,k}, k)$ is a function of event-trigger error, $e_{ET,k}$, and the traditional NN reconstruction error, $\varepsilon(x_k, k)$. An arbitrarily small event-based reconstruction error can be obtained by increasing both the frequency of events and the number of neurons. As a consequence, a properly designed event-trigger condition is necessary. A compromise has to be reached between the reconstruction error and computational load in an event-sampled approximation and control scheme.

3. EVENT BASED OPTIMAL CONTROLLER DESIGN

In this section, the near optimal event-triggered controller design is detailed for the uncertain discrete-time system.

3.1 PROPOSED SOLUTION

The proposed optimal event-triggered control system is illustrated in Figure 1. It consists of: i) a nonlinear discrete-time system, smart sensor and trigger mechanism with a mirror actor-critic network, and ii) an event-based optimal controller. The event-based optimal controller entails three NNs as online approximators: the identifier, critic and actor NNs. These three NNs are used to approximate the system dynamics, time-varying cost function which is the solution to the HJB equation, and the control input, respectively. All the NNs use activation functions with event-sampled inputs. The NN weights are updated in an aperiodic manner at the trigger instants only.

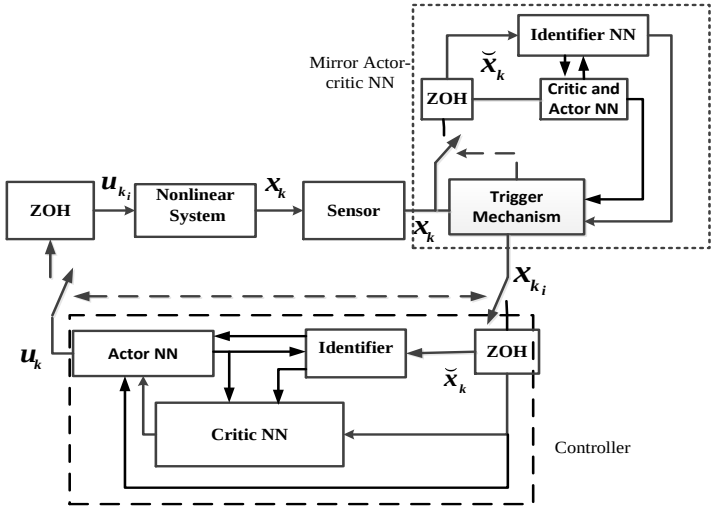

Figure 1. Block diagram representation of the event-triggered control system

The event-trigger instants, k_i for $i = 1, 2, \cdots$ are decided by the smart sensor and the trigger-mechanism. The event-trigger condition is evaluated at every time instant k to determine the trigger instants. At the trigger instants $k = k_i$ for $i = 1, 2, \cdots$, the current system state vector, x_{k_i} , and its previous value $x_{k_{i-1}}$ are together sent to the controller. These event-sampled state vectors are subsequently used to update the NN weights and control input. The updated value of the control input is then sent to the system and held by the ZOH, and utilized until the next update.

Most importantly, the event-trigger condition is made adaptive by designing a suitable threshold. This adaptive trigger condition ensures online approximation of nonlinear functions, as discussed in Remark 1. The threshold is designed as function of the actor NN weight estimates and the system state vector. To evaluate the event-trigger condition, the trigger mechanism consists of a mirror actor-critic NN (see Figure 1). This mirror actor-critic NN operates in synchronism with the one at the controller. Both the actor-critic neural networks are initialized with same initial values. The NN weights are adjusted with events. Thus, the adaptive trigger-condition gets updated at every-trigger instant.

Remark 2: The mirror actor-critic NN estimates the NN weights locally at the trigger mechanism thus relaxing the need for the transmission of NN weights from the controller to the trigger mechanism in the case of NCS. Therefore, the transmission cost only depends upon the transmission of system state and control input vector. Although, the addition of a mirror actor-critic NN increases the computational cost, the overall computation is still reduced due to the event-based execution (also see simulation section).

3.2 IDENTIFIER DESIGN

The input coefficient matrix function $g(x_k)$ is required to compute the optimal control policy (9). This will be generated by the NN based identifier. The universal approximation property of NNs, in a compact set, can be used to represent the nonlinear system in (1). It is given by

$$x_{k+1} = W_I^T \sigma_I(x_k)\bar{u}_k + \varepsilon_I(x_k), \tag{11}$$

where $W_I = [W_f^T \quad W_g^T]^T \in \Re^{(m+1)l_I \times n}$ denotes the unknown constant target weight matrix of the identifier NN. The matrices $W_f \in \Re^{l_I \times n}$, $W_g = [W_{g1}^T \quad W_{g2}^T \quad \cdots \quad W_{gm}^T]^T \in \Re^{ml_I \times n}$ and $W_{gp} \in \Re^{l_I \times n}$ for $p = 1, \cdots, m$. The function $\sigma_I(x_k) = diag\{\sigma_f(x_k),\ \sigma_g(x_k)\} \in \Re^{(m+1)l_I \times (m+1)}$ represents the NN activation function matrix where $\sigma_f(x_k) \in \Re^{l_I}$ and $\sigma_g(x_k) = diag\{\sigma_{g1}(x_k), \sigma_{g2}(x_k), \cdots, \sigma_{gm}(x_k)\} \in \Re^{ml_I \times m}$. The reconstruction error is denoted by $\varepsilon_I(x_k) = \varepsilon_f(x_k) + \varepsilon_g(x_k)u_k \in \Re^n$, where $\varepsilon_f(x_k) \in \Re^n$ and $\varepsilon_g(x_k) \in \Re^{n \times m}$ are the traditional reconstruction errors. The augmented control input is denoted as $\bar{u}_k = [1 \quad u_k^T]^T \in \Re^{m+1}$. The subscript f and g are used to denote the variable for the functions $f(x_k)$ and $g(x_k)$, respectively. The number of neurons in the hidden layer is denoted by l_I. The notation $diag\{\bullet\}$ denotes the matrix formed by the activation function vectors as diagonal blocks and the off diagonals are zero vectors of appropriate dimensions.

Assumption 3[16]: The target weight vector, W_I, the activation function, $\sigma_I(x_k)$, and the traditional reconstruction error, $\varepsilon_I(x_k)$, of the NN are upper bounded such that

$\|W_I\| \le W_{I,M}$, $\|\sigma_I(x_k)\| \le \sigma_{I,M}$ and $\|\varepsilon_I(x_k)\| \le \varepsilon_{I,M}$ where $W_{I,M}$, $\sigma_{I,M}$ and $\varepsilon_{I,M}$ are positive constants.

The control input is updated only at the event-trigger instants and requires the approximated identifier dynamics. Therefore, the event-based identifier dynamics can be represented as

$$\hat{x}_{k+1} = \hat{f}(\breve{x}_k) + \hat{g}(\breve{x}_k)u_k, \quad k_i \le k < k_{i+1},\ i = 1, 2, \cdots, \tag{12}$$

where $\hat{x}_k \in \Re^n$ being the identifier state vector at the time instant k. The functions $\hat{f}(\breve{x}_k) \in \Re^n$ and $\hat{g}(\breve{x}_k) \in \Re^{n\times m}$ represent the approximated identifier dynamics. Note that the identifier structure is based on event-sampled states and held during the inter-event time. This novel event-based structure is selected to reduce additional and redundant computation during the inter-event times.

The identifier dynamics (12) with NN approximation can be written as

$$\hat{x}_{k+1} = \hat{W}_{I,k}^T \sigma_I(\breve{x}_k)\bar{u}_k,\ k_i \le k < k_{i+1}\ , \tag{13}$$

where $\hat{W}_{I,k} = [\hat{W}_{f,k}^T \quad \hat{W}_{g,k}^T]^T \in \Re^{(m+1)l_l \times n}$ is the actual estimated weight matrix, and $\sigma_I(\breve{x}_k) \in \Re^{(m+1)l_l \times (m+1)}$ is the event-sampled activation function matrix for the identifier NN.

The identification error can be defined as $e_{I,k} = x_k - \hat{x}_k$. Hence, the identification error dynamics using this equation with (11) and (13) are found to be

$$e_{I,k+1} = \tilde{W}_{I,k}^T \sigma_I(x_k)\bar{u}_k + \hat{W}_{I,k}^T \left(\sigma_I(x_k) - \sigma_I(\breve{x}_k)\right)\bar{u}_k + \varepsilon_{I,k}, \tag{14}$$

for $k_i \le k < k_{i+1},\ i = 1, 2, \cdots$ where $\tilde{W}_{I.k} = W_I - \hat{W}_{I,k}$ is the identifier NN weight estimation error. The reconstruction error is denoted by $\varepsilon_{I,k} = \varepsilon_I(x_k)$ for brevity.

Consider the case when an event is triggered, i.e., $\breve{x}_k = x_k$ for $k = k_i$. The identifier dynamics in (13) with the updated state vector can be expressed as

$$\hat{x}_{k+1} = \hat{W}_{I,k}^T \sigma_I(x_k)\bar{u}_k,\ k = k_i,\ i = 1,2,\cdots. \tag{15}$$

Therefore, the identification error dynamics from (14) for $k = k_i$ is written as

$$e_{I,k+1} = \tilde{W}_{I,k}^T \sigma_I(x_k)\bar{u}_k + \varepsilon_{I,k},\ k = k_i,\ i = 1,2,\cdots. \tag{16}$$

The aperiodic event-based tuning law for NN identifier weights now can be selected as

$$\hat{W}_{I,k} = \begin{cases} \hat{W}_{I,k-1} + \dfrac{\alpha_I \sigma_I(x_{k-1})\bar{u}_{k-1} e_{I,k}^T}{[\sigma_I(x_{k-1})\bar{u}_{k-1}]^T[\sigma_I(x_{k-1})\bar{u}_{k-1}]+1}, & k = k_i, \\ \hat{W}_{I,k-1}, & k_{i-1} < k < k_i, \end{cases} \tag{17}$$

where α_I is the learning gain. The update law (17) requires the state vector x_{k-1} to compute $e_{I,k}$ at trigger instant $k = k_i$. Hence, the current state, x_k and previous state, x_{k-1} are together sent to the controller once an event is triggered at k, as proposed in Section III.A. The weight update law (17) is aperiodic in nature to save computation.

The identifier NN weight estimation error dynamics from (17), forwarding one time instant ahead, can be expressed as

$$\tilde{W}_{I,k+1} = \begin{cases} \tilde{W}_{I,k} - \dfrac{\alpha_I \sigma_I(x_k)\bar{u}_k e_{I,k+1}^T}{[\sigma_I(x_k)\bar{u}_k]^T[\sigma_I(x_k)\bar{u}_k]+1}, & k = k_i, \\ \tilde{W}_{I,k}, & k_i < k < k_{i+1}. \end{cases} \tag{18}$$

The ultimate boundedness of the identifier NN weight estimation error is guaranteed by the following lemma. Before introducing the lemma, the following assumption is needed.

Assumption 4: The identifier NN activation function $\sigma_I(x_k)$ is Lipschitz continuous in a compact set for all $x_k \in D_x$. Then, there exists a constant C_{σ_I} such that $\left\|\sigma_I(x_k)-\sigma_I(\breve{x}_k)\right\| \le C_{\sigma_i}\left\|x_k-\breve{x}_k\right\|$.

Lemma 1: Consider the nonlinear discrete-time system (1) along with the identifier (13). Assume the Assumptions 1 through 4 hold and the NN initial weights, $\hat{W}_{I,0}$, is initialized in a compact set. Let the identifier NN weights are tuned by (17) at the event-trigger instants and the activation function $\sigma_I(x_k)$ satisfies the persistency of excitation (PE) condition [16]. Suppose the control input is stabilizing and the learning gain α_I satisfy $0<\alpha_I<1/2$. Then, there exist two positive integers T and $\bar{T}$ such that the weight estimation error $\tilde{W}_{I,k}$ is ultimately bounded (UB) with a bound $B_{\tilde{W},I}^{M}$ for all $k_i > k_0 + T$ or, alternatively, $k \ge k_0 + \bar{T}$.

Proof: Refer to the Appendix.

The stabilizing assumption for the control input is later relaxed in the closed-loop stability proof by using an initial admissible control.

3.3 CONTROLLER DESIGN

In this subsection, event-based actor-critic NN designs are presented. Besides the HJB or temporal difference (TD) error, an additional error term corresponding to the terminal cost is defined and used to tune the Critic NN such that the terminal cost can be properly satisfied.

3.3.1 Critic NN Design. Consider the Bellman equation (7). It can be rewritten as

$$0 = V(x_{k+1}, k+1) + Q(x_k, k) + u_k^T R u_k - V(x_k, k). \tag{19}$$

The cost function in (5) using the universal approximation property of NN [16] in a compact set can be written as

$$V(x_k, k) = W_V^T \varphi(x_k, k) + \varepsilon_{V,k}, \tag{20}$$

where $W_V \in \Re^{l_V}$ is the unknown constant target critic NN weights. The time-varying activation function is denoted as $\varphi(x_k, k) \in \Re^{l_V}$. The traditional NN reconstruction error $\varepsilon_{V,k} = \varepsilon_V(x_k, k) \in \Re$, for brevity. The number of hidden layer neurons in the network is given by l_V. The following assumption holds for the critic NN.

Assumption 5[15]: The target NN weights, activation functions and the reconstruction errors of the critic NN are bounded above and satisfy $\|W_V\| \le W_{V,M}$, $\|\varphi(\bullet,\bullet)\| \le \varphi_M$, and $|\varepsilon_V(\bullet,\bullet)| \le \varepsilon_{V,M}$ where $W_{V,M}$ φ_M, and $\varepsilon_{V,M}$ are positive constants. The gradient of the activation function and reconstruction error satisfy $\|\partial\varphi(\bullet,k)/\partial(\bullet)\| \le \nabla\varphi_M$ and $\|\partial\varepsilon_V(\bullet,k)/\partial(\bullet,k)\| \le \nabla\varepsilon_{V,M}$, where $\nabla\varphi_M$ and $\nabla\varepsilon_{V,M}$ are positive constants. In addition, the activation function, $\varphi(x_k, k)$, is Lipschitz continuous for all $x_k \in \mathcal{D}_x$ and satisfies $\|\varphi(x_k, k) - \varphi(\breve{x}_k, k)\| \le C_\varphi \|x_k - \breve{x}_k\| = C_\varphi \|e_{ET,k}\|$ where C_φ is a positive constant.

The Bellman equation (19) using (20) can be expressed as

$$0 = W_V^T \Delta\varphi(x_k, k) + Q(x_k, k) + u_k^T R u_k + \Delta\varepsilon_{V,k}, \tag{21}$$

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

The approximated/estimated cost function by the critic NN with the event-based system states, $\breve{x}_k$, can be represented as

$$\hat{V}\left(\breve{x}_k, k\right) = \hat{W}_{V,k}^T \varphi(\breve{x}_k, k),\ k_i \le k < k_{i+1},\ i = 1, 2, \cdots, \tag{22}$$

where $\hat{W}_{V,k}^T \in \Re^{l_V}$ is the estimated weight, and $\varphi(\breve{x}_k, k) \in \Re^{l_V}$ is the event-based time varying activation function. The activation function is selected such that $\left\|\varphi\left(0, k\right)\right\| = 0$ for $\left\|x_k\right\| = 0$ in order to ensure $\hat{V}\left(0\right) = 0$.

The approximated cost function (22) with the event-based availability of the system state $\breve{x}_k$ for $k_i \le k < k_{i+1},\ i = 1, 2, \cdots$, does not satisfies the relation (21). Therefore, the HJB error or the temporal difference (TD) error, $e_{HJB,k}$, associated with (21) can be written as

$$e_{HJB,k} = Q(\breve{x}_k, k) + u_k^T R u_k + \hat{V}(\breve{x}_{k+1}, k+1) - \hat{V}(\breve{x}_k, k), \tag{23}$$

for $k_i \le k < k_{i+1},\ i = 1, 2, \cdots$. The positive definite function $Q(\breve{x}_k, k)$ is a function of the event-based state vector.

The TD error (23) with the approximated cost function (22) can be represented as

$$e_{HJB,k} = \hat{W}_{V,k}^T \Delta\varphi(\breve{x}_k, k) + Q(\breve{x}_k, k) + u_k^T R u_k,\ k_i \le k < k_{i+1}, \tag{24}$$

where $\hat{V}(\breve{x}_{k+1}, k+1) = \hat{W}_{V,k}^T \varphi(\breve{x}_{k+1}, k+1)$, and $\Delta\varphi(\breve{x}_k, k) = \varphi(\breve{x}_{k+1}, k+1) - \varphi(\breve{x}_k, k)$.

The terminal cost (6) in term of NN approximation(20) can also be represented as

$$V(x_N, N) = W_V^T \varphi(x_N, N) + \varepsilon_{V,N}, \tag{25}$$

where $\varphi(x_N, N)$ and $\varepsilon_{V,N} = \varepsilon_V(x_N, N)$ are the activation function and the reconstruction error, respectively, at the terminal time N.

The approximated/estimated terminal cost from (22) can be expressed as

$$\hat{V}(x_N, N) = \hat{W}_{V,N}^T \varphi(x_N, N). \tag{26}$$

The terminal state vector, x_N, is not known. Thus, it not possible to compute the terminal cost (26) at time k, and hence, the actual terminal cost error. Therefore, a projected terminal cost error, $e_{FC,k}$, can be represented as the difference between the desired terminal cost and the estimated cost at time instant, k. It is represented by

$$e_{FC,k} = \psi(x_N, N) - \hat{W}_{V,k}^T \varphi(\breve{x}_k, N),\ k_i \le k < k_{i+1}, i = 1, 2, \cdots. \tag{27}$$

Note, the activation function, $\varphi(\breve{x}_k, N)$, is an explicit function of final time and the final time N is known. Thus, we can compute $\varphi(\breve{x}_k, N)$ at time k.

The total error in cost function estimation becomes

$$e_{total,k} = e_{HJB,k} + e_{FC,k}, k_i \le k < k_{i+1},\ i = 1, 2, \cdots. \tag{28}$$

At the event-trigger instant, $k = k_i$, $i = 1, 2, \cdots$, the HJB equation or TD error can be written from (24) as

$$e_{HJB,k} = \hat{W}_{V,k}^T \Delta\varphi(x_k, k) + Q(x_k, k) + u_k^T R u_k, \tag{29}$$

where $\Delta\varphi(x_k, k) = \varphi(x_{k+1}, k+1) - \varphi(x_k, k)$. Similarly, the terminal cost error from (27) for $k = k_i, i = 1, 2, \cdots$ becomes

$$e_{FC,k} = \psi(x_N, N) - \hat{W}_{V,k}^T \varphi(x_k, N). \tag{30}$$

The total error at trigger instant by combining (29) and (30) becomes

$$e_{total,k} = \hat{W}_{V,k}^T \Delta\bar{\varphi}(x_k,k) + Q(x_k,k) + u_k^T R u_k + \psi(x_N,N), \tag{31}$$

for $k = k_i, i = 1,2,\cdots$, where $\Delta\bar{\varphi}(x_k,k) = \Delta\varphi(x_k,k) - \varphi(x_k,N)$.

To minimize the total error in an event-triggered context, the critic NN weights are proposed to be updated at the trigger instants for $k = k_i$ and held constant during the inter-event duration, $k_i < k < k_{i+1}$. With this effect, using the previous values for implementation point of view, the update law of critic NN can be selected as

$$\hat{W}_{V,k} = \begin{cases} \hat{W}_{V,k-1} - \dfrac{\alpha_V \Delta\bar{\varphi}(x_{k-1},k-1) e_{total,k-1}^T}{\Delta\bar{\varphi}^T(x_{k-1},k-1)\Delta\bar{\varphi}(x_{k-1},k-1)+1}, & k = k_i \\ \hat{W}_{V,k-1}, & k_{i-1} < k < k_i \end{cases}, \tag{32}$$

where $\alpha_V > 0$ is the learning gain, $\Delta\varphi(x_{k-1},k-1) = \varphi(x_k,k) - \varphi(x_{k-1},k-1)$. The total error $e_{total,k-1}$ cab be computed from (31) by moving one time step backward.

Remark 3: Similar to the identifier NN, the critic NN weights are updated in an aperiodic manner. This further saves computation when compared to traditional NN based control.

Adding the difference between (24) and (21) to (27), the total error can be represented in terms of the critic NN weight estimation error, $\tilde{W}_{V,k} = W_V - \hat{W}_{V,k}$. It is found to be

$$\begin{aligned} e_{\text{total},k} &= -\tilde{W}_{V,k}^T \Delta\bar{\varphi}(\bar{x}_k,k) - W_V^T \Delta\tilde{\varphi}(x_k,\bar{x}_k,k) - \tilde{Q}(x_k,\bar{x}_k,k) \\ &+ W_V^T(\varphi(x_N,N) - \varphi(\bar{x}_k,N)) - \Delta\bar{\varepsilon}_{V,k}, \quad k_i \le k < k_{i+1}, \end{aligned} \tag{33}$$

where $\Delta\bar{\varepsilon}_{V,k} = \Delta\varepsilon_{V,k} - \varepsilon_{V,N}$, $\Delta\bar{\varphi}(\breve{x}_k, k) = \Delta\varphi(\breve{x}_k, k) - \varphi(\breve{x}_k, N)$, $\tilde{Q}(x_k, \breve{x}_k, k) = Q(x_k, k) - Q(\breve{x}_k, k)$, $\Delta\tilde{\varphi}(x_k, \breve{x}_k, k) = \Delta\varphi(x_k, k) - \Delta\varphi(\breve{x}_k, k)$. It is routine to check $\left\|\Delta\bar{\varphi}(\bullet,\bullet)\right\| \le \Delta\bar{\varphi}_M$ and $\left\|\Delta\bar{\varepsilon}_{V,k}\right\| \le \Delta\bar{\varepsilon}_{V,M}$ from Assumption 5. The total error at the event-trigger instant from (33) with $\breve{x}_k = x_k$ for $k = k_i$ becomes

$$e_{total,k} = -\tilde{W}_{V,k}^T \Delta\bar{\varphi}(x_k, k) + W_V^T(\varphi(x_N, N) - \varphi(x_k, N)) - \Delta\bar{\varepsilon}_{V,k}. \tag{34}$$

The critic NN weight estimation error dynamics, from (32) by moving one step forward, can be expressed as

$$\tilde{W}_{V,k+1} = \begin{cases} \tilde{W}_{V,k} + \dfrac{\alpha_V \Delta\bar{\varphi}(x_k, k) e_{total,k}^T}{\Delta\bar{\varphi}^T(x_k, k)\Delta\bar{\varphi}(x_k, k) + 1}, & k = k_i \\ \tilde{W}_{V,k}, & k_i < k < k_{i+1} \end{cases}, \tag{35}$$

The last task is to design the actor NN which is given next.

3.3.2 Actor Design. In this subsection, we approximate the optimal control policy through the actor NN to implement it forward in time. The identified control coefficient matrix of the NN identifier is also used to update the actor NN.

The optimal control input (9) by the approximation property of NN [16] in a compact set can be written as

$$u_k^* = W_u^T \sigma_u(x_k, k) + \varepsilon_{u,k}, \tag{36}$$

where $W_u \in \Re^{l_u \times m}$ is the unknown constant target weight matrix. The time varying activation function is denoted by $\sigma_u(x_k, k) \in \Re^{l_u}$ and the traditional reconstruction error is $\varepsilon_{u,k} = \varepsilon_u(x_k, k) \in \Re^m$. The number of neurons of the actor NN is given by l_u.

Assumption 6: The target NN weights, activation function, and the reconstruction error of the actor NN are upper bounded and satisfy $\left\|W_u\right\| \le W_{u,M}$, $\left\|\sigma_u(\bullet,\bullet)\right\| \le \sigma_{u,M}$ and

$\left\|\varepsilon_u(\bullet,\bullet)\right\| \le \varepsilon_{u,M}$, where $W_{u,M}$, $\sigma_{u,M}$, and $\varepsilon_{u,M}$ are positive constants. The actor NN activation function is Lipschitz continuous for all $x_k \in \mathcal{D}_x$ such that $\left\|\sigma_u(x_k,k)-\sigma_u(\breve{x}_k,k)\right\| \le C_{\sigma_u}\left\|x_k-\breve{x}_k\right\| = C_{\sigma_u}\left\|e_{ET,k}\right\|$ where C_{σ_u} is a positive constant.

Moreover, the optimal control input (9) by using gradient of cost function (20) can be expressed as

$$u_{V,k}^{*} = -(1/2)R^{-1}g^T(x_k)\nabla\varphi^T(x_{k+1},k+1)W_V - (1/2)R^{-1}g^T(x_k)\nabla\varepsilon_{V,k+1}, \tag{37}$$

where $\nabla\varphi(x_{k+1},k+1) = \partial\varphi(x_{k+1},k+1)/\partial x_{k+1}$ and $\nabla\varepsilon_{V,k+1} = \partial\varepsilon_V(x_{k+1})/\partial x_{k+1}$. Both the optimal control inputs (36) and (37)should be equal. Their difference can be expressed as

$$0 = W_u^T\sigma_u(x_k,k) + \varepsilon_{u,k} + (1/2)R^{-1}g^T(x_k)\nabla\varphi^T(x_{k+1},k+1)W_V + (1/2)R^{-1}g^T(x_k)\nabla\varepsilon_{V,k+1}. \tag{38}$$

The approximated/estimated optimal control input by the actor NN in an event-trigger context can be represented as

$$u_k = \hat{W}_{u,k}^T\sigma_u(\breve{x}_k,k),\ k_i \le k < k_{i+1},\ i = 1,2,\cdots \tag{39}$$

where $\hat{W}_{u,k} \in \Re^{l_u \times m}$ is the estimated actor NN weights, and $\sigma_u(\breve{x}_k,k) \in \Re^{l_u}$ denotes the time varying event-based activation function.

Further, the estimated control input, $u_{V,k}$, using the gradient of the estimated cost function (22), can also be written as

$$u_{V,k} = -(1/2)R^{-1}\hat{g}^T(\breve{x}_k)\nabla\varphi^T(\breve{x}_{k+1},k+1)\hat{W}_{V,k}, \tag{40}$$

for $k_i \le k < k_{i+1}$, $i = 1,2,\cdots$, where $\hat{g}(\breve{x})$ is the approximated event-based control coefficient matrix from the NN-based identifier and $\nabla\varphi(\breve{x}_{k+1},k+1) = \partial\varphi(\breve{x}_{k+1},k+1)/\partial x_{k+1}$. The control policy (39) applied to the system (1) and the control policy (40) which

minimizes the estimated cost function (22) will not satisfy (38). Hence, the control input estimation error, $e_{u,k}$ for $k_i \le k < k_{i+1},\ i = 1, 2, \cdots$ is represented as the difference between (39) and (40), and found to be

$$e_{u,k} = \hat{W}_{u,k}^T \sigma_u(\breve{x}_k, k) + (1/2) R^{-1} \hat{g}^T(\breve{x}_k) \nabla \varphi^T(\breve{x}_{k+1}, k+1) \hat{W}_{V,k}. \tag{41}$$

Similar to the critic NN, the actor NN weights are proposed to be updated at the trigger instants only. The update law of actor NN, using previous values, is chosen as

$$\hat{W}_{u,k} = \begin{cases} \hat{W}_{u,k-1} - \dfrac{\alpha_u \sigma_u(x_{k-1}, k-1) e_{u,k-1}^T}{\sigma_u^T(x_{k-1}, k-1)\sigma_u(x_{k-1}, k-1) + 1}, & k = k_i, \\ \hat{W}_{u,k-1}, & k_{i-1} < k < k_i, \end{cases} \tag{42}$$

where α_u is the learning gain. The control input estimation error $e_{u,k-1}$ can be computed form (41) by moving one time step backward at the trigger instant and given by

$$e_{u,k-1} = \hat{W}_{u,k-1}^T \sigma_u(\breve{x}_{k-1}, k-1) + (1/2) R^{-1} \hat{g}^T(\breve{x}_{k-1}) \nabla \varphi^T(x_{k-1}, k-1) \hat{W}_{V,k-1}. \tag{43}$$

The control input estimation error can be expressed in terms of the actor NN weight estimation error, $\tilde{W}_{u,k}$, by subtracting (38) from (41). This is described by

$$\begin{aligned} e_{u,k} = {} & -\tilde{W}_{u,k}^T \sigma_u(\breve{x}_k, k) - (1/2) R^{-1} g^T(\breve{x}_k) \nabla \varphi^T(\breve{x}_{k+1}, k+1) \tilde{W}_{V,k} \\ & + (1/2) R^{-1} \tilde{g}^T(x_k) \nabla \varphi^T(\breve{x}_{k+1}, k+1) \tilde{W}_{V,k} - (1/2) R^{-1} \tilde{g}^T(\breve{x}_k) \nabla \varphi^T(\breve{x}_{k+1}, k+1) W_V \\ & + \varepsilon_{u,k}^{sum1}, \quad k_i \le k < k_{i+1}, i = 1, 2, \cdots \end{aligned} \tag{44}$$

where $\varepsilon_{u,k}^{sum1} = -W_u^T \sigma_u(x_k, \breve{x}_k, k) - (1/2) R^{-1} g^T(\breve{x}_k) \nabla \tilde{\varphi}^T(x_{k+1}, \breve{x}_{k+1}, k+1) W_V - (1/2) R^{-1} \tilde{g}^T(x_k, \breve{x}_k)$

$\times \nabla \varphi^T(x_{k+1}, k+1) W_V - (1/2) R^{-1} g^T(x_k) \nabla \varepsilon_{V,k+1} - \varepsilon_{u,k}$ with $\tilde{g}(x_k, \breve{x}_k) = g(x_k) - g(\breve{x}_k)$,

$\sigma_u(x_k, \breve{x}_k, k) = \sigma_u(x_k, k) - \sigma_u(\breve{x}_k, k)$ and $\nabla \tilde{\varphi}(x_{k+1}, \breve{x}_{k+1}, k+1) = \nabla \varphi(x_{k+1}, k+1) - \nabla \varphi(\breve{x}_{k+1}, k+1)$.

It clear that $0 \le \left\| \varepsilon_{u,k}^{sum1} \right\| \le \varepsilon_{u,M}^{sum1}$ where $\varepsilon_{u,M}^{sum1}$ is a positive constant.

Further, from (44) the control input estimation error at $k = k_i, i = 1,2,\cdots$ can be written as

$$\begin{aligned} e_{u,k} &= -\tilde{W}_{u,k}^T \sigma_u(x_k,k) - (1/2) R^{-1} g^T(x_k) \nabla \varphi^T(x_{k+1},k+1) \tilde{W}_{V,k} + (1/2) R^{-1} \tilde{g}^T(x_k) \\ &\times \nabla \varphi^T(x_{k+1},k+1) \tilde{W}_{V,k} - (1/2) R^{-1} \tilde{g}^T(x_k) \nabla \varphi^T(x_{k+1},k+1) W_V + \varepsilon_{u,k}^{sum}, \end{aligned} \tag{45}$$

where $\varepsilon_{u,k}^{sum} = -(1/2) R^{-1} g^T(x_k) \nabla \varepsilon_{V,k+1} - \varepsilon_{u,k}$ and $\left\| \varepsilon_{u,k}^{sum} \right\| \le \varepsilon_{u,M}^{sum}$.

The weight estimation error dynamics of the actor NN, from (42), moving one time step ahead, becomes

$$\tilde{W}_{u,k+1} = \begin{cases} \tilde{W}_{u,k} + \dfrac{\alpha_u \sigma_u(x_k,k) e_{u,k}^T}{\sigma_u^T(x_k,k)\sigma_u(x_k,k)+1}, & k = k_i, \\ \tilde{W}_{u,k}, & k_i < k < k_{i+1}. \end{cases} \tag{46}$$

Next, the main results of the near optimal event-triggered system are claimed.

4. EVENT TRIGGER CONDITION AND STABILITY ANALYSIS

In this section we formulate the closed-loop event-triggered dynamics. The main results are claimed by designing an adaptive event-trigger condition. The closed-loop system dynamics are obtained by using (1) , the actual control input (39) , the ideal control input (36) with simple mathematical manipulation and given by

$$\begin{aligned} x_{k+1} &= f(x_k) + g(x_k)u_k^* - g(x_k)\left(\tilde{W}_{u,k}^T\sigma_u(x_k,k) + \varepsilon_{u,k}\right) - g(x_k)\hat{W}_{u,k}^T \\ &\times\left(\sigma_u(x_k,k) - \sigma_u(\breve{x}_k,k)\right),\ k_i \le k < k_{i+1}. \end{aligned} \tag{47}$$

At the event-trigger instants, $k = k_i$ with updated state vector, the closed-loop system dynamics can be rewritten from (47) as

$$x_{k+1} = f(x_k) + g(x_k)u_k^* - g(x_k)\left(\tilde{W}_{u,k}^T\sigma_u(x_k,k) + \varepsilon_{u,k}\right). \tag{48}$$

Before, claiming the main result in the theorem, the following lemma is necessary.

Lemma 2[15]: Consider the nonlinear discrete-time system given by (1). Then, there exist an optimal control policy u_k^* for (1) such that the closed-loop dynamics satisfies the inequality

$$\left\| f(x_k) + g(x_k)u_k^* \right\|^2 \le K^* \left\| x_k \right\|^2, \tag{49}$$

where $0 < K^* < 1$ is a constant.

Now consider the event-trigger error (2). The following condition is selected as the event-trigger condition:

$$D\left(\left\| e_{ET,k} \right\|\right) \le \sigma_{ET,k} \left\| x_k \right\|, \tag{50}$$

where the threshold coefficient is denoted by

$$\sigma_{ET,k} = \sqrt{(1-2K^*)\Gamma_{ET} \Big/ 4g_M^2 C_{\sigma_u}^2 \left\| \hat{W}_{u,k} \right\|^2}, \tag{51}$$

with $0<\Gamma_{ET}<1$, $0<K^{*}<1/2$. The dead-zone operator $D(\bullet)$ is defined as

$$D\left(\left\|e_{ET,k}\right\|\right)=\begin{cases}\left\|e_{ET,k}\right\|, & \left\|x_k\right\|>b_x,\\ 0, & \text{otherwise,}\end{cases} \tag{52}$$

with b_x being the ultimate bound for the state. The system state and the control input vectors are transmitted to the controller and the plant, respectively, when the event-trigger condition in (50) is not satisfied (or violated). Further, an event is trigger when the estimated NN weight $\left\|\hat{w}_{u,k}\right\|=0$ to update the NN weights without evaluating the trigger condition.

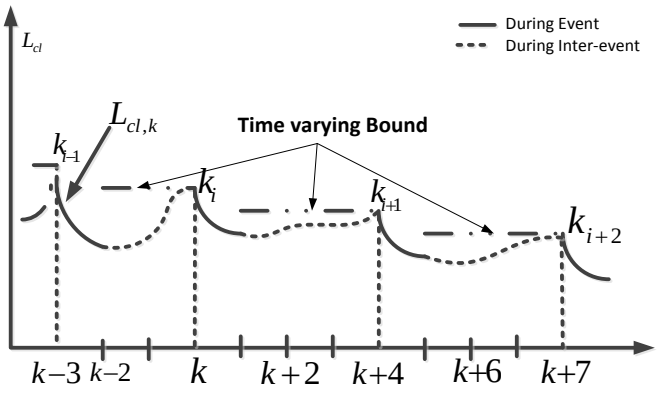

Figure 2. Evolution of the Lyapunov function.

Next, the theorem guarantees the UB of the closed-loop event-trigger system. The UB is shown by using a Lyapunov function for both cases of triggering, i.e., at the events and inter-events. It is important to mention that, the Lyapunov function is not monotonically converging to the ultimate bound both during both the events and inter-event times. This is also not necessary to show stability of the system as discussed in [4] for event-trigged control system, and **[**17**]** for switched systems. Therefore, in our case,

during the inter-event times the Lyapunov function is allowed to increase but within a time varying upper bound. Further, it is shown that with trigger of events the time varying upper bound and the Lyapunov function converge to the UB as illustrated in the Figure 2.

Theorem 2: Consider the nonlinear discrete-time system (1), the NN identifier (13), NN critic (22) and NN actor networks (39). Assume $u_0(x_k)$ be an initial stabilizing control policy for the system (1) and Assumptions 1 through 6 hold. Let the identifier, critic and actor NN weight estimates , $\hat{W}_{I,0}$, $\hat{W}_{V,0}$ and $\hat{W}_{u,0}$, respectively, are initialized in their respective compact sets. Suppose, the system state vector is sent to the controller and the NN weights are updated using (17), (32) and (42) through the violation of the event-trigger condition (50). Let the activation functions $\sigma_I(x_k)$, $\varphi(x_k,k)$ and $\sigma_u(x_k,k)$ satisfy the PE condition [16]. Then, there exists positive constants $0<\alpha_I<1/2$, $0<\alpha_V<1/3$ and $0<\alpha_u<1/5$ such that the closed-loop event-triggered system state vector, x_k , the identifier, critic and actor NN weight estimation errors $\tilde{W}_{I,k}$, $\tilde{W}_{V,k}$ and $\tilde{W}_{u,k}$, respectively, are UB for all $k_i \geq k_0 + T$ or, alternatively $k \geq k_0 + \bar{T}$. Further, $\left\|V^* - \hat{V}\right\| \leq b_V$ and $\left\|u^* - u\right\| \leq b_u$ with b_V and b_u are small positive constants.

Proof: Refer to the Appendix.

Remark 4: The selection of $0<K^*<1/2$ satisfies Lemma 2 and varies according to the desired performance of the system. The adaptive event-trigger condition (50) with (51) implicitly depends upon the actor NN weight estimation error, $\tilde{W}_{u,k}$. During the initial learning phase, the NN weight estimation error will be large. Hence, the events are

triggered frequently. This facilitates the approximation of the cost function, control policy and the system dynamics to achieve near optimal performance.

Remark 5: The dead zone operator (52) used with the event-trigger condition helps to stop unnecessary triggering due to the NN reconstruction error. The dead zone is enabled once the system state is in the ultimate bound b_x. The bound $b_x = \max\left(b_{1,M}^x, b_{2,M}^x\right)$ computed from (A.14) and (A.18) is a function of the tuning parameters α_I, α_V, α_u and the NN reconstruction error bounds $\varepsilon_{I,M}$, $\varepsilon_{V,M}$, and $\varepsilon_{u,M}$. Therefore, the bound can be made arbitrarily small as mentioned in the Remark A.1**.**

5. NON-TRIVIAL MINIMUM INTER-EVENT TIME

In this section we discuss the non-triviality of the inter event times for the near optimal event-triggered control system. An explicit formula for the minimum inter-event time is also presented. The minimum inter-event time is the minimum time interval between two consecutive event-sampling instants over all sampling instants, i.e., $\delta k_{\min} = \min_{i\in\mathbb{N}}\{\delta k_i\}$ where $\delta k_i = k_{i+1} - k_i$ for $i = 1, 2, \cdots$ are the inter-event times. This is implicitly defined by the event-trigger condition (50). In case of a discrete time system, the minimum inter-event time is trivial and becomes the sampling time, T_s or $\delta k_{\min} = 1$. So, it is important to guarantee nontrivial inter-event times, i.e., $\delta k_i > 1$ to reduce the computational load. In the case of approximation-based control design, the inter-event times largely depend on NN weight estimation error and presented in the following theorem.

Theorem 3: Let the hypothesis in Theorem 2 holds. The minimum inter event-time can be expressed as

$$\delta k_{\min} = \min_{i\in\mathbb{N}}\{\ln\left(1+(1/N_i)((M_i-1)\sigma_{ET,\min})\right)/\ln(M_i)\}, \tag{53}$$

for $i = 1, 2, \cdots$ and the non-triviality of the inter-event times are guaranteed if the following condition is satisfied:

$$\ln(1+(1/N_i)(M_i-1)\sigma_{ET,\min}) > \ln(M_i), i = 1, 2, \cdots \tag{54}$$

where $N_i = ((\sqrt{K^*}+1)\|x_{k_i}\| + g_M(\sigma_{u,M}\|\tilde{W}_{u,k_i}\| + \varepsilon_{u,M}))$ and $M_i = (\sqrt{K^*} + g_M C_{\sigma_u}\|\hat{W}_{u,k_i}\|)$ and $\sigma_{ET,\min} = \min_{k\in\mathbb{N}}\{\sigma_{ET,k}\|x_k\|\}$ is the minimum event-trigger threshold.

Proof: Refer to Appendix.

Remark 6: It is important to note that the inter-event times will be non-trivial, i.e., $\delta k_i > 1,\ i = 1, 2, \cdots$ if (54) is satisfied. To achieve nontrivial inter trigger times during the initial learning, the initial NN weights needs to be selected close to the target parameters. This will reduce the NN weight estimation error, $\tilde{W}_{u,k}$ which in turn decreases the value of N_i and increases M_i in (53). Thus, the condition (54) is satisfied leading to non-trivial inter-event times. In addition, along with the update of the NN weights, $\hat{W}_{u,k}$, the weight estimation error, $\tilde{W}_{u,k}$, will further decrease and, hence, the variable N_i. This, further, ensures elongated inter-event times.

The proposed event sampled design can be used mutatis mutandis for nonlinear networked control systems (NNCS) in the presence of time varying network induced delays and random packet losses. The detailed design procedure along with the simulation results are presented in Appendix B of the dissertation. The NNCS is represented as a continuous time nonlinear system in affine form. The proposed event sampled ADP design discussed in Section 3 through 5 is extended for a stochastic design due to random network constraints. An infinite horizon cost function is minimized to design the event sampled control policy. It was observed that the event sampled stochastic ADP for NNCS resulted in 66% saving in computational load and 27% saving in the network bandwidth usage.

6. SIMULATION RESULTS

In this section, a two-link robot has been considered for simulation. The continuous time dynamics of the two-link robot is given by (1)

$$\dot{x} = f(x) + g(x)u, \tag{55}$$

with internal dynamics $f(x_k)$ and control coefficient matrix $g(x_k)$ given as

where $f(x) = \begin{bmatrix} x_3 \\ x_4 \\ \dfrac{\begin{pmatrix} -(2x_3x_4 + x_4^2 - x_3^2 - x_3^2\cos x_2)\sin x_2 \\ +20\cos x_1 - 10\cos(x_1 + x_2)\cos x_2 \end{pmatrix}}{\cos^2 x_2 - 2} \\ \dfrac{\begin{pmatrix} (2x_3x_4 + x_4^2 + 2x_3x_4\cos x_2 + x_4^2\cos x_2 + 3x_3^2 \\ +2x_3^2\cos x_2 + 20(\cos(x_1 + x_2) - \cos x_1)\times \\ (1+\cos x_2) - 10\cos x_2\cos(x_1 + x_2) \end{pmatrix}}{\cos^2 x_2 - 2} \end{bmatrix}$ and $g(x) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ \dfrac{1}{2-\cos^2 x_2} & \dfrac{-1-\cos x_2}{2-\cos^2 x_2} \\ \dfrac{-1-\cos x_2}{2-\cos^2 x_2} & \dfrac{3+2\cos x_2}{2-\cos^2 x_2} \end{bmatrix}$

The continuous dynamics was discretized first for simulation. The following simulation parameters were selected to carry out the simulation. The cost-to-go was selected as quadratic function with $Q(x_k) = x_k^T Q_x x_k$, $Q_x = I_{4\times4}$ and $R = 0.001 * I_{2\times2}$ where I is the identity matrix. The non-quadratic terminal cost was chosen as $\psi(x_N, N) = 1$. The initial weights for the critic NN were selected as zero. The actor and the identifier NN weights were initialized with random values from a uniform distribution in the interval of zero to one. The time-varying activation functions for both the critic and actor NNs were constructed as state-dependent and time-dependent terms, i.e., $\varphi(x_k, k) = \varphi_t(k)\varphi_x(x_k)$. The state-dependent part, $\varphi_x(x_k)$, was chosen as $\varphi_x(x_k) = \{x_{1,k}^2, \cdots, x_{4,k}^2, x_{1,k}x_{2,k}, \cdots, x_{1,k}^3 x_{2,k}, \cdots, x_{1,k}^4, \cdots, x_{4,k}^4, \cdots, x_{1,k}x_{2,k}x_{3,k}x_{4,k}\} \in \Re^{45\times1}$ [18] and the time-dependent part, $\varphi_t(k)$, was also selected as

$\varphi_t(k) = \{1, [\exp(-\tau)]^1, ..., [\exp(-\tau)]^{44}; \cdots; [\exp(-\tau)]^{44}, [\exp(-\tau)]^{43}, ..., 1\} \in \Re^{45\times 45}$ [15] where $\tau = (\mathrm{N} - k)/\mathrm{N}$ is the normalized time index. The identifier activation function was chosen as $\tanh\{(x_{1,k})^2, x_{1,k}x_{2,k}, \cdots, (x_{1,k})^5(x_{2,k}), ..., (x_{4,k})^6\}$. The number of neurons for the identifier was 39 and the critic and the action NN were 45 each.

The learning rates for the NN tuning were selected as $\alpha_I = 0.03$, $\alpha_V = 0.01$ and $\alpha_u = 0.05$ per the conditions derived in Theorem 2. The event-trigger condition parameters were $K^* = 0.45$, $\Gamma_{ET} = 0.92$, $C_{\sigma_u} = 2$ and $g_M = 1.5$. The initial admissible control was selected as $u_0 = [-500x_1 - 500x_3,\ -200x_2 - 200x_4]^T$ and the terminal time was $N = 10000$. The ultimate bound selected for the system state was 0.0005. The event-trigger threshold was computed using (50), with (51), and (52) with the above parameters selected for simulation.

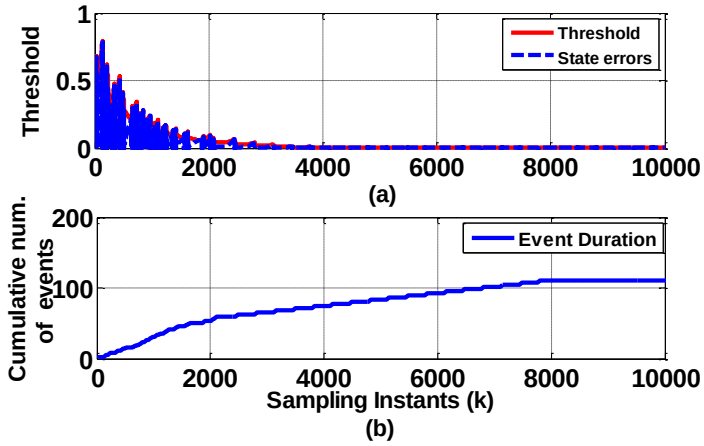

Figure 3. (a) Triggering threshold with event-trigger error; (b) cumulative number of triggered events vs. sampling instants.

Figure 3 (a) shows the evolution of the threshold (solid) over time along with the event-trigger error (dotted). From this figure it is evident that the event-trigger error reset to zero once it reaches the threshold with trigger of events. In Figure 3 (b), the cumulative number of trigger instants is plotted against the total sampling instants. Even though a large number of triggering occurs in the initial phase, the cumulative number of triggers is reduced. The cumulative triggering became constant after 8000 time instants. This implies the system state is in the ultimate bound $b_x = 5\times10^{-4}$. The number of events during the sampling time of 10 sec with a sampling interval of 0.001 sec was found to be 110.

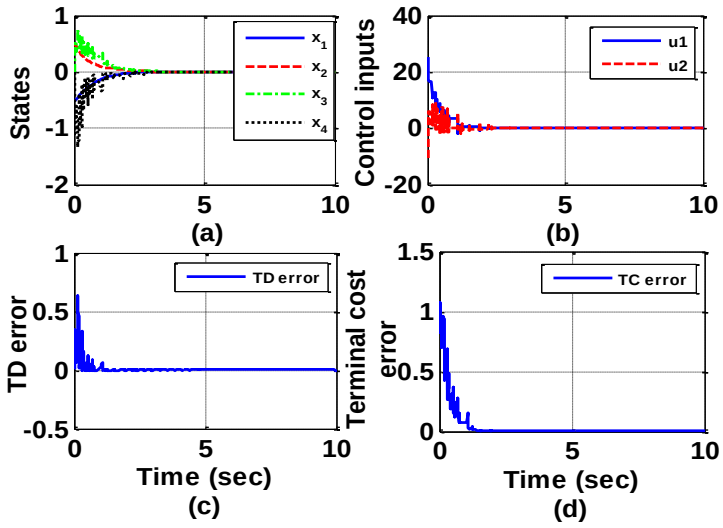

Figure 4. Convergence of (a) system state; (b) near optimal control inputs; (c) HJB error; and (d) terminal cost error.

A comparison of the computational load in terms of the multiplication and addition that is required to compute the event-trigger condition and controller is given in Table 1. It indicates a reduction in computation of around 65.5% for the event-triggered

system. Furthermore, if a communication network is included between the plant and the controller, fewer transmissions are needed due to event based sampling. This will reduce the communication cost significantly.

The performance of the optimal controller is shown in Figure 4. The optimal control input (Figure 4 (b)) regulates the system states to zero as shown in Figure 4(a). The control input also converges to zero with system states. This implies that with a reduced number of controller executions the system is near optimally regulated. Further, the HJB equation error, shown in Figure 4 (c), converges to near zero implying the optimality achieved in finite time. The terminal cost error also converges to near zero and shown in Figure 4 (d).

Table 1. Comparison of computational load between traditional and event-based discrete time systems.

Systems		Number of additions and Multiplications	Sampling instants	Total computation
Traditional discrete time	Identifier	10	10000	310000
	Critic	13		
	Actor	8		
Event-based discrete time	Identifier and mirror	20	110	106820
	Critic and mirror	26		
	Actor and mirror	16		
	Trig. Con (periodic)	10		

7. CONCLUSIONS

In this paper, a near optimal event-triggered control of an uncertain nonlinear discrete time system in affine form is introduced. The actor-critic frame work used to solve the finite horizon optimal control problem with event-based approximation was able to regulate the system. The novel adaptive event-trigger condition generated the required number of events at the initial learning phase to achieve a small approximation error. This also saved the computation by fewer updates in the control law. Near optimality was achieved in a finite time with complete unknown system dynamics. With an explicit formula, it is shown that a nontrivial inter-event time can exist with proper initialization of weights and event-based NN weight updates. It was observed that the cumulative number of triggered events varies with initial NN weights. The effectiveness of the controller is validated using simulation.

8. REFERENCES

[1] K. J. Åström and B. M. Bernhardsson, "Comparison of periodic and event based sampling for the first order stochastic systems," in *Proceedings of the 14th IFAC World Congress*, Beijing, China, Jul. 1999, pp. 301–306.

[2] 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.

[3] E. Garcia and P. J. Antsaklis, "Model-based event-triggered control for systems with quantization and time-varying network delays," *IEEE Transactions on Automatic Control*, vol. 58, no. 2, pp. 422–434, Feb. 2013.

[4] X. Wang and M. D. Lemmon, "On event design in event-triggered feedback systems," *Automatica*, vol. 47, no. 10, pp. 2319-2322, Oct. 2011.

[5] W. Heemels and M. C. F. Donkers, "Model-based periodic event-triggered control for linear systems" *Automatica*, vol. 49, no. 3 , pp. 698-711, Mar. 2013.

[6] 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 *Proceedings of the American Control Conference*, Wasington, DC, USA, Jun. 2013, pp. 6433-6438.

[7] A. Sahoo, H. Xu, and S. Jagannathan, "Neural network-based adaptive event-triggered control of nonlinear continuous-time systems," in *Proceedings of the IEEE Multi-conference on Systems and Control*, Hyderabad, India, Aug. 2013, pp. 35-40.

[8] X. Hao, S. Jagannathan and F. L. Lewis, "Stochastic optimal control of unknown networked control systems in the presence of random delays and packet losses," *Automatica*, vol. 48, no. 6, pp. 1017–1030, Jun. 2012.

[9] F. L. Lewis and V. L. Syrmos, *Optimal Control*. Hoboken, NY: John Wiley and Sons, 1995.

[10] O.C. Imer and T. Basar, "To measure or to control: optimal control with scheduled measurements and controls," in *Proceedings of the American Control Conference*, Minneapolis, MN, USA, Jun. 2006, pp. 14-16.

[11] A. Molin and S. Hirche, "On the optimality of certainty equivalence for event-triggered control systems," *IEEE Transactions on Automatic Control*, vol. 58, no. 2 , pp. 470-474, Jun. 2013.

[12] M. Rabi, K. H. Johansson, and M. Johansson, "Optimal stopping for event-triggered sensing and actuation," in *Proceedings of the* 47th *IEEE Conference on Decision and Control*, Cancun, Mexico, Dec. 2008, pp.3607-3612.

[13] D. P. Bertsekas and J. N. Tsitsiklis, *Neuro*-dynamic *Programming*. Belmonth, MA: Athena Scientific, 1996.

[14] H. Xu and S. Jagannathan, "Neural network based finite horizon stochastic optimal controller design for nonlinear networked control systems," in *Procedinngs of the IEEE International Joint Conference on Neural Network*, Dallas, TX, USA, Aug. 2013 pp. 1-7.

[15] 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 Transactions on Neural Network and Learning Systems*, vol. 23, no. 7 , pp. 1118-1129, May 2012.

[16] S. Jagannathan, *Neural Network Control of Nonlinear Discrete-time Systems*. Boca Raton, FL: CRC Press, 2006.

[17] H. Ye, A. Michel, and L. Hou, "Stability theory for hybrid dynamical systems," *IEEE Transactions on Automatic Control*, vol. 43, no. 4, pp. 461–474, Apr. 1998.

[18] Z. Chen and S. Jagannathan, "Generalized Hamilton-Jacobi-Bellman formulation-based neural netwrok control of affine nonlinear discrete-time system," *IEEE Transaction on Neural Networks*, vol. 19, no. 1 , pp. 90-106, Jan. 2008.

[19] G.Cybenko,"Approximation by super positions of sigmoidal functions," *Mathematics of Control Signals and Systems*, vol. 2, no. 4, pp. 304-314, Dec. 1989.

[20] G. Bitsoris and E. Gravalou, "Comparison principle, positive invariance and constrained regulation of nonlinear systems," *Automatica*, vol. 31, no. 2 , pp. 217–222, Feb. 1995.

APPENDIX

Proof of Theorem 1: The smooth and continuous function $h(x_k, k)$, with the universal approximation theorem [19] of NN, can be represented in a compact set as

$$h(x_k, k) = W^T \sigma(x_k, k) + \varepsilon(x_k, k), \tag{A.1}$$

with x_k as input to the activation function at every sampling instant k. Consider the event-based sampling where the state x_k is available intermittently as defined in (3). Equation (A.1) can be expressed as

$$h(x_k, k) = W^T \sigma(x_k, k) - W^T \sigma(\breve{x}_k, k) + W^T \sigma(\breve{x}_k, k) + \varepsilon(x_k, k), \tag{A.2}$$

where $\sigma(\breve{x}_k, k)$ and $\breve{x}$ are is event-based activation function and state vectors.

The state, x_k, in terms of the event-based state, $\breve{x}_k$, and event-trigger error, $e_{ET,k}$, in (2) can be written as $x_k = \vartheta(\breve{x}_k, e_{ET,k}) = \breve{x}_k + e_{ET,k}$. Substituting this expression, (A.2) can be represented as

$$h(x_k, k) = W^T \sigma(\breve{x}_k, k) + \varepsilon_e(\breve{x}_k, e_{ET,k}, k), \tag{A.3}$$

where $\varepsilon_e(\breve{x}_k, e_{ET,k}, k) = W^T[\sigma(\vartheta(\breve{x}_k, e_{ET,k}), k) - \sigma(\breve{x}_k, k)] + \varepsilon(\vartheta(\breve{x}_k, e_{ET,k}), k)$. ■

Proof of Lemma 1: The UB of the identifier weight estimation error is proven by demonstrating the boundedness of the weight estimation error for both trigger conditions, separately. A single Lyapunov function is used to evaluate the first difference and combined at the end to show overall UB.

Case I : *Event Triggered, i.e.,* $k = k_i, i = 1, 2, \cdots$

Consider a Lyapunov function candidate given by

$$L_{I,k} = tr\{\tilde{W}_{I,k}^T \tilde{W}_{I,k}\}. \tag{A.4}$$

The first difference, $\Delta L_{I,k} = tr\{\tilde{W}_{I,k+1}^T \tilde{W}_{I,k+1}\} - tr\{\tilde{W}_{I,k}^T \tilde{W}_{I,k}\}$, along the dynamics of the identifier NN weight estimation error (18) for $k = k_i$, becomes

$$\Delta L_{I,k} = -2\alpha_I tr\left\{\tilde{W}_{I,k}^T \sigma_I(x_k)\bar{u}_k e_{I,k+1}^T\right\} / ([\sigma_I(x_k)\bar{u}_k]^T[\sigma_I(x_k)\bar{u}_k]+1)$$
$$+\alpha_I^2 tr\left\{e_{I,k+1}[\sigma_I(x_k)\bar{u}_k]^T[\sigma_I(x_k)\bar{u}_k]e_{I,k+1}^T\right\} / ([\sigma_I(x_k)\bar{u}_k]^T[\sigma_I(x_k)\bar{u}_k]+1)^2 .$$

Substituting the identification error dynamics (16), and using Cauchy-Schwartz (C-S) inequality with the fact that $[\sigma_I(x_k)\bar{u}_k]^T[\sigma_I(x_k)\bar{u}_k] / ([\sigma_I(x_k)\bar{u}_k]^T[\sigma_I(x_k)\bar{u}_k]+1) \le 1$, the first difference is bounded by

$$\Delta L_{I,k} \le -\alpha_I(1-2\alpha_I) tr\left\{\frac{\tilde{W}_{I,k}^T[\sigma_I(x_k)\bar{u}_k][\sigma_I(x_k)\bar{u}_k]^T\tilde{W}_{I,k}}{[\sigma_I(x_k)\bar{u}_k]^T[\sigma_I(x_k)\bar{u}_k]+1}\right\} + \frac{\alpha_I(2+\alpha_I)}{\left\|\sigma_I(x_k)\bar{u}_k\right\|^2+1}\left\|\varepsilon_{I,k}\right\|^2 .$$

By definition, the augmented control input $\left\|\bar{u}_k\right\| \ge 1$, and, $0 < \sigma_{I,\mathrm{m}} \le \left\|\sigma_I(x_k)\right\| \le \sigma_{I,M}$ is satisfied due to the PE condition [15] and Assumption 3. Hence, $0 < \sigma_{I,m} \le \left\|\sigma_I(x_k)\bar{u}_k\right\|$. By the above facts, the first term in the above equation satisfy

$$tr\left\{\frac{\tilde{W}_{I,k}^T[\sigma_I(x_k)\bar{u}_k][\sigma_I(x_k)\bar{u}_k]^T\tilde{W}_{I,k}}{[\sigma_I(x_k)\bar{u}_k]^T[\sigma_I(x_k)\bar{u}_k]+1}\right\} = \frac{\left\|\tilde{W}_{I,k}^T\sigma_I(x_k)\bar{u}_k\right\|^2}{(\left\|\sigma_I(x_k)\bar{u}_k\right\|^2+1)} \ge \frac{\left\|\tilde{W}_{I,k}^T\sigma_I(x_k)\bar{u}_k\right\|^2 / \left\|\bar{u}_k\right\|^2}{(\left\|\sigma_I(x_k)\right\|^2\left\|\bar{u}_k\right\|^2+1)/\left\|\bar{u}_k\right\|^2}$$

$$\ge \frac{\Omega_{I,m}^2}{\sigma_{I,M}^2+1}\left\|\tilde{W}_{I,k}\right\|^2 \quad \text{where} \quad 0 < \Omega_{I,m} \le \left\|\sigma_I(x_k)\bar{u}_k / \left\|\bar{u}_k\right\|\right\| \text{ with a bounded control input.}$$

Substituting the above inequality:

$$\Delta L_{I,k} \le -\alpha_I(1-2\alpha_I)\left(\Omega_{I,m}^2 / (\sigma_{I,M}^2+1)\right)\left\|\tilde{W}_{I,k}\right\|^2 + B_{\tilde{W}_I}, \tag{A.5}$$

where $B_{\tilde{W}_I} = \alpha_I(1+2\alpha_I)\varepsilon_{I,M}^2 / (1+\sigma_{I,m}^2)$. From (A.5), by selecting $0 < \alpha_I < 1/2$, the Lyapunov first difference $\Delta L_{I,k} < 0$ as long as

$$\left\|\tilde{W}_{I,k}\right\| > \sqrt{(\sigma_{I,M}^2+1)B_{\tilde{W}_I} / \alpha_I(1-2\alpha_I)\Omega_{I,m}^2} = B_{\tilde{W},I}^M .$$

Therefore, by the Lyapunov theorem [16], the identifier weight estimation error, $\tilde{W}_{I,k}$ is UB with a bound $B_{\tilde{W},I}^{M}$ for all $k_i \geq k_0 + T$ with the occurrence of events.

Case II: *Event not triggered, i.e.,* $k_i < k < k_{i+1}$.

Consider the same Lyapunov function (A.4). The first difference along the identifier weight estimation error dynamics (18) for $k_i < k < k_{i+1}$

$$\Delta L_{I,k} = tr\{\tilde{W}_{I,k+1}^{T}\tilde{W}_{I,k+1}\} - tr\{\tilde{W}_{I,k}^{T}\tilde{W}_{I,k}\} = 0\,. \tag{A.6}$$

From (A.6) the Lyapunov first difference, $\Delta L_{I,k}$, during the inter-event time remains at zero. This implies the NN weight estimation error, $\tilde{W}_{I,k}$, remains constant during the inter-events times. The initial weight estimate, $\hat{W}_{I,0}$, is finite and from Assumption 3 the target weight matrix is bounded. Therefore, initial weight estimation error, $\tilde{W}_{I,0}$ is also bounded. Further, $\tilde{W}_{I,k}$ is bounded at the trigger instants as shown in Case I. Thus, the initial value $\tilde{W}_{I,k_i}$, $i = 1, 2, \cdots$, for each inter-event time, which is the updated value at the previous trigger instant, is also bounded. Consequently, the weight estimation error, $\tilde{W}_{I,k}$, is constant and bounded during the inter-event times, i.e., $k_i < k < k_{i+1}$ for $i = 1, 2, \cdots$.

From Cases I and II, the identifier weight estimation error is bounded both at the trigger instants and inter-event times. Further, with the occurrence of events followed by each inter-event time, the identifier weight estimation error, $\tilde{W}_{I,k}$, is UB with a bound $B_{\tilde{W},I}^{M}$ for all $k_i \geq k_0 + T$. Alternatively, $\tilde{W}_{I,k}$ is UB for all $k \geq k_0 + \bar{T}$ as k_i is a subsequence of k and $\bar{T}$ is a function of T. ■

Proof of Theorem 2: The stability of the closed-loop system is proved by considering both the event-conditions, i.e., event is triggered, $k = k_i$ and event is not triggered, $k_i < k < k_{i+1}$, $i = 1, 2, \cdots$. A single Lyapunov function is evaluated for both cases, separately, and combined at the end to show the convergence to the UB.

Case 1: *Event triggered, i.e.,* $k = k_i, i = 1, 2, \cdots$.

Consider the Lyapunov function candidate given by

$$L_{cl,k} = L_{x,k} + L_{I,k} + L_{V,k} + L_{u,k} + + L_{A,k} + L_{B,k}, \tag{A.7}$$

where $L_{x,k} = \Gamma_x x_k^T x_k$, $L_{I,k} = \Gamma_I tr\{\tilde{W}_{I,k}^T \tilde{W}_{I,k}\}$, $L_{V,k} = \Gamma_V \tilde{W}_{V,k}^T \tilde{W}_{V,k}$ $L_{u,k} = tr\left\{\tilde{W}_{u,k}^T \tilde{W}_{u,k}\right\}$, $L_{A,k} = \Gamma_{I2} tr\{\tilde{W}_{I,k}^T \tilde{W}_{I,k}\}^2$, and $L_{B,k} = \Gamma_{V2}\{\tilde{W}_{V,k}^T \tilde{W}_{V,k}\}^2$. The positive constants

$$\Gamma_x = \frac{\alpha_u (1-5\alpha_u)\sigma_{u,m}^2}{8 g_M^2 \sigma_{u,M}^2 (\sigma_{u,M}^2 + 1)}, \qquad \Gamma_I = \frac{2\varpi\left(\sigma_{I,M}^2 + 1\right)}{\alpha_I (1-2\alpha_I)\Omega_{I,m}^2}, \qquad \Gamma_V = \frac{2\vartheta\left(\Delta\bar{\varphi}_M^2 + 1\right)}{\alpha_V (1-3\alpha_V)\Delta\bar{\varphi}_m^2},$$

$$\Gamma_{I2} = \frac{2\alpha_u (4+5\alpha_u)\lambda_{\max}^2\left(R^{-1}\right)\nabla\varphi_M^2 \sigma_{I,M}^2 \left(\sigma_{I,M}^2 + 1\right)^2}{4\alpha_I (1-2\alpha_I)\left(2(\sigma_{I,M}^2 + 1) - \alpha_I (1-2\alpha_I)\Omega_{I,m}^2\right)\Omega_{I,m}^2 (\sigma_{u,m}^2 + 1)},$$

$$\Gamma_{V2} = \frac{2\alpha_u (4+5\alpha_u)\lambda_{\max}^2\left(R^{-1}\right)\nabla\varphi_M^2 \sigma_{I,M}^2 \left(\Delta\bar{\varphi}_M^2 + 1\right)^2}{4\alpha_V (1-3\alpha_V)\left(2\left(\Delta\bar{\varphi}_M^2 + 1\right) - \alpha_V (1-3\alpha_V)\Delta\bar{\varphi}_m^2\right)\Delta\bar{\varphi}_m^2 (\sigma_{u,m}^2 + 1)},$$

$$\varpi = \frac{\alpha_u (4+5\alpha_u)\lambda_{\max}^2\left(R^{-1}\right)\nabla\varphi_M^2 W_{V,M}^2 \sigma_{I,M}^2}{2\left(\sigma_{u,m}^2 + 1\right)} + \frac{\Gamma_{I2} B_{\tilde{W}_I}\left(2\left(\sigma_{I,M}^2 + 1\right) - \alpha_I (1-2\alpha_I)\Omega_{I,m}^2\right)}{\left(\sigma_{I,M}^2 + 1\right)},$$

and

$$\vartheta = \frac{\alpha_u (4+5\alpha_u)\lambda_{\max}^2\left(R^{-1}\right)\left(g_M^2 + 2\varepsilon_{I,M}^2\right)\nabla\varphi_M^2}{4(\sigma_{u,m}^2 + 1)} + \frac{\Gamma_{V2}\varepsilon_V^{1,M}\left(2(\Delta\bar{\varphi}_M^2 + 1) - \alpha_V (1-3\alpha_V)\Delta\bar{\varphi}_m^2\right)}{(\Delta\bar{\varphi}_M^2 + 1)}.$$

Consider the first term in the Lyapunov function candidate (A.7), $L_{x,k} = \Gamma_x x_k^T x_k$. The first difference along the closed-loop system dynamics (48) is bounded above by

$$\Delta L_{x,k} \le \Gamma_x \left\| x_{k+1} \right\|^2 - \Gamma_x \left\| x_k \right\|^2$$
$$\le \Gamma_x \left\| f(x_k) + g(x_k) u_k^* - g(x_k)(\tilde{W}_{u,k}^T \sigma_u(x_k) + \varepsilon_{u,k}) \right\|^2 - \Gamma_x \left\| x_k \right\|^2 .$$

Recalling the Lemma 2 and applying C-S inequality $(a^2 + b^2) \le 2a^2 + 2b^2$, it reveals that

$$\Delta L_{x,k} \le -(1 - 2K^*)\Gamma_x \left\| x_k \right\|^2 + 4 g_M^2 \sigma_{u,M}^2 \Gamma_x \left\| \tilde{W}_{u,k} \right\|^2 + 4\Gamma_x g_M^2 \varepsilon_{u,M}^2 . \tag{A.8}$$

Consider the second term in the Lyapunov function (A.7), $L_{I,k} = \Gamma_I tr\{\tilde{W}_{I,k}^T \tilde{W}_{I,k}\}$. The first difference can be written from (A.5) and is given by

$$\Delta L_{I,k} \le -\Gamma_I \alpha_I (1 - 2\alpha_I) \left(\Omega_{I,m}^2 / (\sigma_{I,M}^2 + 1) \right) \left\| \tilde{W}_{I,k} \right\|^2 + \Gamma_I B_{\tilde{W}_I} . \tag{A.9}$$

Moving on for the third term $L_{V,k} = \Gamma_V \tilde{W}_{V,k}^T \tilde{W}_{V,k}$, in the Lyapunov function candidate (A.7), the first difference becomes $\Delta L_{V,k} = \tilde{W}_{V,k+1}^T \tilde{W}_{V,k+1} - \tilde{W}_{V,k}^T \tilde{W}_{V,k}$. Along the critic NN weight estimation error dynamics (35) for $k = k_i$, the first difference can be represented as

$$\Delta L_{V,k} = \frac{2\Gamma_V \alpha_V \tilde{W}_{V,k}^T \Delta\bar{\varphi}(x_k,k) e_{total,k}^T}{\Delta\bar{\varphi}^T(x_k,k)\Delta\bar{\varphi}(x_k,k)+1} + \frac{\Gamma_V \alpha_V^2 e_{total,k} \Delta\bar{\varphi}^T(x_k,k)\Delta\bar{\varphi}(x_k,k) e_{total,k}^T}{(\Delta\bar{\varphi}^T(x_k,k)\Delta\bar{\varphi}(x_k,k)+1)^2} .$$

Substituting (34) into the above equation and using the C-S inequality, the first difference of Lyapunov function candidate

$$\Delta L_{V,k} \le -\frac{2\Gamma_V \alpha_V \tilde{W}_{V,k}^T \Delta\bar{\varphi}(x_k,k)\Delta\bar{\varphi}^T(x_k,k)\tilde{W}_{V,k}}{\Delta\bar{\varphi}^T(x_k,k)\Delta\bar{\varphi}(x_k,k)+1}$$
$$+ \frac{2\Gamma_V \alpha_V \tilde{W}_{V,k}^T \Delta\bar{\varphi}(x_k,k)\left(\varphi(x_N,N) - \varphi(x_k,N)\right)^T W_V^T}{\Delta\bar{\varphi}^T(x_k,k)\Delta\bar{\varphi}(x_k,k)+1} - \frac{2\Gamma_V \alpha_V \tilde{W}_{V,k}^T \Delta\bar{\varphi}(x_k,k)\Delta\bar{\varepsilon}_{V,k}^T}{\Delta\bar{\varphi}^T(x_k,k)\Delta\bar{\varphi}(x_k,k)+1}$$
$$+ \frac{3\Gamma_V \alpha_V^2 \tilde{W}_{V,k}^T \Delta\bar{\varphi}(x_k,k)\Delta\bar{\varphi}^T(x_k,k)\Delta\bar{\varphi}(x_k,k)\Delta\bar{\varphi}^T(x_k,k)\tilde{W}_{V,k}}{\left(\Delta\bar{\varphi}^T(x_k,k)\Delta\bar{\varphi}(x_k,k)+1\right)^2}$$

$$+\frac{3\Gamma_V\alpha_V^2 W_V^T\left(\varphi(x_N,N)-\varphi(x_k,N)\right)\Delta\bar{\varphi}^T(x_k,k)\Delta\bar{\varphi}(x_k,k)\left(\varphi(x_N,N)-\varphi(x_k,N)\right)^T W_V}{\left(\Delta\bar{\varphi}^T(x_k,k)\Delta\bar{\varphi}(x_k,k)+1\right)^2}$$

$$+\frac{3\Gamma_V\alpha_V^2\Delta\bar{\varepsilon}_{V,k}\Delta\bar{\varphi}^T(x_k,k)\Delta\bar{\varphi}(x_k,k)\Delta\bar{\varepsilon}_{V,k}^T}{\left(\Delta\bar{\varphi}^T(x_k,k)\Delta\bar{\varphi}(x_k,k)+1\right)^2}.$$

By using Young's inequality $2a^Tb\le qa^Ta+(1/q)b^Tb$, with $q>0$, $\Delta\bar{\varphi}^T(x_k,k)\Delta\bar{\varphi}(x_k,k)/(\Delta\bar{\varphi}^T(x_k,k)\Delta\bar{\varphi}(x_k,k)+1)<1$ and $1/\left(\Delta\bar{\varphi}^T(x_k,k)\Delta\bar{\varphi}(x_k,k)+1\right)\le 1$, the first difference becomes

$$\Delta L_{V,k}\le-\Gamma_V\alpha_V\left(1-3\alpha_V\right)\frac{\tilde{W}_{V,k}^T\Delta\bar{\varphi}(x_k,k)\Delta\bar{\varphi}^T(x_k,k)\tilde{W}_{V,k}}{\Delta\bar{\varphi}^T(x_k,k)\Delta\bar{\varphi}(x_k,k)+1}$$

$$+\frac{\Gamma_V\alpha_V\left(2+3\alpha_V\right)W_V^T\left(\varphi(x_N,N)-\varphi(x_k,N)\right)\left(\varphi(x_N,N)-\varphi(x_k,N)\right)^T W_V}{\Delta\bar{\varphi}^T(x_k,k)\Delta\bar{\varphi}(x_k,k)+1}$$

$$+\frac{\Gamma_V\alpha_V\left(2+3\alpha_V\right)\Delta\bar{\varepsilon}_{V,k}\Delta\bar{\varepsilon}_{V,k}^T}{\Delta\bar{\varphi}^T(x_k,k)\Delta\bar{\varphi}(x_k,k)+1}.$$

From Assumption 5, $\varphi(x_N,N)-\varphi(x_k,k)<2\varphi_M$ and $\left\|\Delta\bar{\varepsilon}_{V,k}\right\|\le\Delta\bar{\varepsilon}_{V,M}$. With these facts and simple manipulation using C-S inequality and Frobenius norm, we arrive at

$$\Delta L_{V,k}\le-\Gamma_V\alpha_V\left(1-3\alpha_V\right)\left(\Delta\bar{\varphi}_m^2/(\Delta\bar{\varphi}_M^2+1)\right)\left\|\tilde{W}_{V,k}\right\|^2+\Gamma_V\varepsilon_V^{1,M}, \tag{A.10}$$

where $\varepsilon_V^{1,M}=\alpha_V\left(2+3\alpha_V\right)\left(\varphi_M^2W_{V,M}^2/(\Delta\bar{\varphi}_m^2+1)\right)+\alpha_V\left(2+3\alpha_V\right)(\Delta\bar{\varepsilon}_{V,M}^2/\Delta\bar{\varphi}_m^2+1)$, $0<\alpha_V<1/3$ and $0<\Delta\bar{\varphi}_{\mathrm{m}}\le\left\|\Delta\bar{\varphi}(x_k,k)\right\|\le\Delta\bar{\varphi}_M$ which is satisfied by ensuring PE condition [15].

Consider the next term in the Lyapunov function candidate (A.7), $L_{u,k}=tr\left\{\tilde{W}_{u,k}^T\tilde{W}_{u,k}\right\}$. The first difference along the actor NN weight estimation error dynamics (46) for $k=k_i$ becomes

$$\Delta L_{u,k}=2\alpha_u tr\left\{\frac{\tilde{W}_{u,k}^T\sigma_u(x_k,k)e_{u,k}^T}{(\sigma_u^T(x_k,k)\sigma_u(x_k,k)+1)}\right\}+\alpha_u^2 tr\left\{\frac{e_{u,k}\sigma_u^T(x_k,k)\sigma_u(x_k,k)e_{u,k}^T}{\left(\sigma_u^T(x_k,k)\sigma_u(x_k,k)+1\right)^2}\right\}.$$

Substitute the control input estimation error $e_{u,k}$ from (45) in the above equation. After some mathematical manipulation using C-S inequality and the fact $\sigma_u^T(x_k,k)\sigma_u(x_k,k)/(\sigma_u^T(x_k,k)\sigma_u(x_k,k)+1)<1$, we arrive at

$$\begin{aligned}
\Delta L_{u,k} \le & -\alpha_u\left(1-5\alpha_u\right)tr\{\frac{\tilde{W}_{u,k}^T\sigma_u(x_k,k)\sigma_u^T(x_k,k)\tilde{W}_{u,k}}{\sigma_u^T(x_k,k)\sigma_u(x_k,k)+1}\} \\
& +\alpha_u\left(4+5\alpha_u\right)tr\{\frac{(R^{-1}g^T(x_k)\nabla\varphi^T(x_{k+1},k+1)\tilde{W}_{V,k})(R^{-1}g^T(x_k)\nabla\varphi^T(x_{k+1},k+1)\tilde{W}_{V,k})^T}{4(\sigma_u^T(x_k,k)\sigma_u(x_k,k)+1)}\} \\
& +\alpha_u\left(4+5\alpha_u\right)tr\{\frac{(R^{-1}\tilde{g}^T(x_k)\nabla\varphi^T(x_{k+1},k+1)\tilde{W}_{V,k})}{4(\sigma_u^T(x_k,k)\sigma_u(x_k,k)+1)}(R^{-1}\tilde{g}^T(x_k)\nabla\varphi^T(x_{k+1},k+1)\tilde{W}_{V,k})^T\} \\
& +\alpha_u\left(4+5\alpha_u\right)tr\{\frac{(R^{-1}\tilde{g}^T(x_k)\nabla\varphi^T(x_{k+1},k+1)W_V)}{4(\sigma_u^T(x_k,k)\sigma_u(x_k,k)+1)}(R^{-1}\tilde{g}^T(x_k)\nabla\varphi^T(x_{k+1},k+1)W_V)^T\} \\
& +\alpha_u\left(4+5\alpha_u\right)tr\{\varepsilon_{u,k}^{sum})(\varepsilon_{u,k}^{sum})^T/(\sigma_u^T(x_k,k)\sigma_u(x_k,k)+1)\}.
\end{aligned}$$

Using Frobenius norm, Young's inequality and the relation $\|\tilde{g}(x_k)\| \le \sigma_{I,M}\left\|\tilde{W}_{I,\mathrm{k}}\right\| + \varepsilon_{I,M}$, it holds that

$$\begin{aligned}
\Delta L_{u,k} \le & -\left(\frac{\alpha_u(1-5\alpha_u)\sigma_{u,m}^2}{(\sigma_{u,M}^2+1)}\right)\left\|\tilde{W}_{u,k}\right\|^2+\frac{\alpha_u(4+5\alpha_u)\lambda_{\max}^2(R^{-1})(g_M^2+2\varepsilon_{I,M}^2)\nabla\varphi_M^2}{4(\sigma_{u,m}^2+1)}\left\|\tilde{W}_{V,k}\right\|^2 \\
& +\frac{\alpha_u\left(4+5\alpha_u\right)\lambda_{\max}^2(R^{-1})\sigma_{I,M}^2\nabla\varphi_M^2}{4(\sigma_{u,m}^2+1)}\left\|\tilde{W}_{I,\mathrm{k}}\right\|^4+\frac{\alpha_u\left(4+5\alpha_u\right)\lambda_{\max}^2(R^{-1})\sigma_{I,M}^2\nabla\varphi_M^2}{4(\sigma_{u,m}^2+1)}\left\|\tilde{W}_{V,k}\right\|^4 \\
& +\alpha_u\left(4+5\alpha_u\right)\frac{\lambda_{\max}^2(R^{-1})\sigma_{I,M}^2W_{V,M}^2\nabla\varphi_M^2}{2(\sigma_{u,m}^2+1)}\left\|\tilde{W}_{I,\mathrm{k}}\right\|^2+\alpha_u\left(4+5\alpha_u\right)\frac{\lambda_{\max}^2(R^{-1})W_{V,M}^2\nabla\varphi_M^2\varepsilon_{I,M}^2}{2(\sigma_{u,m}^2+1)} \\
& +\frac{\alpha_u\left(4+5\alpha_u\right)(\varepsilon_{u,M}^{sum})^2}{4(\sigma_{u,m}^2+1)},
\end{aligned} \tag{A.11}$$

where $0<\sigma_{u,m}\le\left\|\sigma_u\left(x_k,k\right)\right\|\le\sigma_{u,M}$, is ensured by the PE condition, $\lambda_{\max}\left(R^{-1}\right)$ is the maximum eigenvalue of R^{-1}.

Considering the next term $L_{A,k}=\Gamma_{I2}tr\left\{\tilde{W}_{I,k}^T\tilde{W}_{I,k}\right\}^2$. The first difference, $\Delta L_{A,k}=\Gamma_{I2}tr\left\{\tilde{W}_{I,k+1}^T\tilde{W}_{I,k+1}\right\}^2-\Gamma_{I2}tr\left\{\tilde{W}_{I,k}^T\tilde{W}_{I,k}\right\}^2$, from (A.9), becomes

$$\Delta L_{A,k} \le -\Gamma_{I2}\alpha_I(1-2\alpha_I)\left(2-\left(\alpha_I(1-2\alpha_I)\Omega_{I,m}^2/(\sigma_{I,M}^2+1)\right)\right)\left(\Omega_{I,m}^2/(\sigma_{I,M}^2+1)\right)\left\|\tilde{W}_{I,k}\right\|^4$$
$$+\Gamma_{I2}\left(B_{\tilde{W}_I}\right)^2+\Gamma_{I2}B_{\tilde{W}_I}\left(2-\left(\alpha_I(1-2\alpha_I)\Omega_{I,m}^2/(\sigma_{I,M}^2+1)\right)\right)\left\|\tilde{W}_{I,k}\right\|^2. \quad \text{(A.12)}$$

where $\left(2-\left(\alpha_I(1-2\alpha_I)\Omega_{I,m}^2/(\sigma_{I,M}^2+1)\right)\right)>0$ with $0<\alpha_u<1/5$.

Similarly, the first difference of the last term $L_{B,k}=\left\{\tilde{W}_{V,k}^T\tilde{W}_{V,k}\right\}^2$ using (A.10) can be written as

$$\Delta L_{B,k} \le -\Gamma_{V2}\alpha_V\left(1-3\alpha_V\right)\left(\Delta\bar{\varphi}_m^2/(\Delta\bar{\varphi}_M^2+1)\right)\left(2-\left(\alpha_V\left(1-3\alpha_V\right)\left(\Delta\bar{\varphi}_m^2/(\Delta\bar{\varphi}_M^2+1)\right)\right)\right)$$
$$\times\left\|\tilde{W}_{V,k}\right\|^4+\Gamma_{V2}\varepsilon_V^{1,M}\left(2-\left(\alpha_V\left(1-3\alpha_V\right)\left(\Delta\bar{\varphi}_m^2/(\Delta\bar{\varphi}_M^2+1)\right)\right)\right)\left\|\tilde{W}_{V,k}\right\|^2+\Gamma_{V2}\left(\varepsilon_V^{1,M}\right)^2. \quad \text{(A.13)}$$

At the final step, combine the individual first differences (A.8), (A.9), (A.10), (A.11), (A.12) and (A.13) to get the overall first difference. Substituting the constants Γ_x, Γ_I, Γ_V, Γ_{I2} and Γ_{V2}, from (A.7), the overall first difference satisfies

$$\Delta L_{cl,k} \le -(1-2K^*)\Gamma_x\left\|x_k\right\|^2-\varpi\left\|\tilde{W}_{I,k}\right\|^2-\vartheta\left\|\tilde{W}_{V,k}\right\|^2-\frac{1}{2}\alpha_u\left(1-5\alpha_u\right)\left(\sigma_{u,m}^2/(\sigma_{u,M}^2+1)\right)$$
$$\times\left\|\tilde{W}_{u,k}\right\|^2-\alpha_u\left(4+5\alpha_u\right)\left(\lambda_{\max}^2\left(R^{-1}\right)\nabla\varphi_M^2\sigma_{I,M}^2/4(\sigma_{u,m}^2+1)\right)\left\|\tilde{W}_{I,\mathrm{k}}\right\|^4+\varepsilon_{cl,total}^{c1} \quad \text{(A.14)}$$
$$-\alpha_u\left(4+5\alpha_u\right)\left(\lambda_{\max}^2\left(R^{-1}\right)\nabla\varphi_M^2\sigma_{I,M}^2/4(\sigma_{u,m}^2+1)\right)\left\|\tilde{W}_{V,k}\right\|^4,$$

where

$$\varepsilon_{cl,total}^{c1}=\Gamma_I B_{\tilde{W}_I}+\Gamma_V\varepsilon_V^{1,M}+\Gamma_{I2}(B_{\tilde{W}_I})^2+(2\alpha_u(4+5\alpha_u)\lambda_{\max}^2(R^{-1})\nabla\varphi_M^2W_{V,M}^2\varepsilon_{I,M}^2/4(\sigma_{u,m}^2+1))$$
$$+(\alpha_u(4+5\alpha_u)(\varepsilon_{u,M}^{sum})^2/(\sigma_{u,m}^2+1))+4\Gamma_x g_M^2\varepsilon_{u,M}^2+\Gamma_{V2}(\varepsilon_V^{1,M})^2.$$

From (A.14), and selecting $0<\alpha_u<1/5$, the first difference of the Lyapunov function, $\Delta L_{cl,k}<0$ as long as

$$\left\|x_k\right\|>\sqrt{\varepsilon_{cl,total}^{c1}/(1-K^*)\Gamma_x}=b_{1,M}^x,$$

$$\left\|\tilde{W}_{I,k}\right\|>\max\{\sqrt[4]{4\left(\sigma_{u,m}^2+1\right)\varepsilon_{cl,total}^{c1}/\alpha_u\left(4+5\alpha_u\right)\lambda_{\max}^2\left(R^{-1}\right)\nabla\varphi_M^2\sigma_{I,M}^2},\sqrt{\varepsilon_{cl,total}^{c1}/\varpi}\}\equiv b_{\tilde{W}_I} \text{ or}$$

$$\left\|\tilde{W}_{V,k}\right\| > \{\sqrt[4]{4\left(\sigma_{u,m}^2+1\right)\varepsilon_{cl,total}^{c1}/\alpha_u\left(4+5\alpha_u\right)\lambda_{\max}^2\left(R^{-1}\right)\nabla\varphi_M^2\sigma_{I,M}^2)},\ \sqrt{\varepsilon_{cl,total}^{c1}/\vartheta}\} \equiv b_{\tilde{W}_V} \text{ or}$$

$$\left\|\tilde{W}_{u,k}\right\| > \sqrt{(2(\sigma_{u,M}^2+1)\varepsilon_{cl,total}^{c1}/\alpha_u\ \overline{(1-5\alpha_u)\sigma_{u,m}^2)} \equiv b_{\tilde{W}_u}}.$$

This implies the system state, x_k the NN weight estimation errors for the identifier, critic and the actor, $\tilde{W}_{I,k}$, $\tilde{W}_{V,k}$, and $\tilde{W}_{u,k}$ are UB for all $k_i \geq k_0 + T$.

Case 2: *Event not triggered, i.e.,* $k_i < k < k_{i+1}, i = 1, 2, \cdots$.

Consider the same Lyapunov function candidate (A.7) as in Case I. The first difference $\Delta L_{x,k} = \Gamma_x \left\|x_{k+1}\right\|^2 - \Gamma_x \left\|x_k\right\|^2$ of the first term along (47) with Lemma 2 and C-S inequality, can be written as

$$\Delta L_{x,k} \leq 2K^* \left\|x_k\right\|^2 + 4g_M^2 \left\|\hat{W}_{u,k}^T \sigma(x_k,k) - \hat{W}_{u,k}^T \sigma(\bar{x}_k,k)\right\|^2 - \left\|x_k\right\|^2 + 4g_M^2(\sigma_{u,M}^2 \left\|\tilde{W}_{u,k}\right\|^2 + \varepsilon_{u,M}^2).$$

From the Lipschitz continuity of the actor NN activation function, in Assumption 6, it holds that

$$\Delta L_{x,k} \leq -(1-2K^*)\left\|x_k\right\|^2 + 4g_M^2 C_{\sigma_u}^2 \left\|\hat{W}_{u,k}\right\|^2 \left\|e_{ET,k}\right\|^2 + \varepsilon_{cl,total,k}^{c2}\ , \tag{A.15}$$

with $\varepsilon_{cl,total,k}^{c2} = 4g_M^2(\sigma_{u,M}^2 \left\|\tilde{W}_{u,k}\right\|^2 + \varepsilon_{u,M}^2)$. Recalling the event-trigger condition (50) for the case when system state vector is outside the ultimate bound, the first difference satisfies

$$\Delta L_{x,k} \leq -(1-2K^*)(1-\Gamma_{ET})\left\|x_k\right\|^2 + \varepsilon_{cl,total,k}^{c2}, \tag{A.16}$$

where $0 < \Gamma_{ET} < 1$ and $0 < K^* < 1/2$.

Considering the remaining terms of Lyapunov function candidates (A.7), the first differences becomes zero due to no update. They are represented as

$$\Delta L_{I,k} = 0,\ \Delta L_{V,k} = 0,\ \Delta L_{A,k} = 0 \text{ and } \Delta L_{B,k} = 0. \tag{A.17}$$

Finally, combining (A.16) and (A.17) the first difference of the overall system is given by

$$\Delta L_{cl,k} \leq -(1-2K^*)(1-\Gamma_{ET})\left\|x_k\right\|^2 + \varepsilon^{c2}_{cl,total,k}\,. \tag{A.18}$$

From (A.18), the first difference $\Delta L_{cl,k} < 0$ as long as

$$\left\|x_k\right\| > \sqrt{\varepsilon^{c2}_{cl,total,k} \big/ (1-2K^*)(1-\Gamma_{ET})} = b^x_{2,k}\,.$$

The actor NN weight estimation error, $\tilde{W}_{u,k}$, is constant during each i^{th} inter-event time, $k_i < k < k_{i+1}$, as the weights are held. Therefore, $\varepsilon^{c2}_{cl,total,k}$ and, hence, $b^x_{2,k}$ are piecewise constant functions. Thus, the system state is bounded by a time varying bound $b^x_{2,k}$ during the inter-event times. The boundedness of the NN weight estimation errors during inter event times can be shown as follows. The NN initial weight estimates are finite. Therefore, the initial the weight estimation errors are also bounded. From Case I, the NN weight estimation errors are bounded at the trigger instants. Therefore, the initial values during inter-event times are bounded. Further, from (A.17), the NN weight estimation errors are remain constant at their respective previous values during the inter-event times. Therefore, the NN weight estimation errors $\tilde{W}_{I,k}$, $\tilde{W}_{V,k}$ and $\tilde{W}_{u,k}$ remain bounded during the inter-event times.

Note that, from Case I, with trigger of events, the system state vector and the NN weight estimation errors converge to UB for all $k_i \geq k_0 + T$. During the inter-event times, from Case II, the system states are bounded by the time varying bound, $b^x_{2,k}$, and NN weight estimation errors are held at their previous values. During the initial learning phase, the piecewise constant bound $b^x_{2,k}$ may be large. Therefore, the system state vector may increase. Alternatively, the Lyapunov function $L_{cl,k}$ may increase during inter-event times, $k_i < k < k_{i+1}$, for $i = 1, 2, \cdots$ as shown in Figure 2. Since the change in system state

vector is governed by the event-trigger condition, a large value of system state vector will lead to an event. Hence, the NN weights and control inputs will be updated which will make the state and weight estimation error to converge.

Further, since each inter-event is followed by an event, the function $\varepsilon_{cl,total,k}^{c2}$ for $k_i < k < k_{i+1}$, in (A.18), is less than the previous inter-event time $k_{i-1} < k < k_i$ and, hence, $b_{2,k}^{x}$. This implies that for all $k_i \geq k_0 + T$, the function $\varepsilon_{cl,total,k}^{c2} \to \varepsilon_{cl,M}^{c2}$ where $\varepsilon_{cl,M}^{c2}$ $= 4g_M^2(\sigma_{u,M}^2 b_{\tilde{W}_u}{}^2 + \varepsilon_{u,M}^2)$ is a constant and $b_{\tilde{W}_u}$ is the ultimate bound for $\tilde{W}_{u,k}$ from Case I. Therefore, the bound for the system state $b_{2,k}^{x}$ will also converge, i.e., $b_{2,k}^{x} \to b_{2,M}^{x}$ for $k_i \geq k_0 + T$ where $b_{2,M}^{x} = \sqrt{\varepsilon_{cl,M}^{c2} / (1-2K^*)(1-\Gamma_{ET})}$ is a small constant.

Consequently, from Case I and Case II, the system state, x_k the NN weight estimation errors for the identifier, critic and the actor, $\tilde{W}_{I,k}$, $\tilde{W}_{V,k}$, and $\tilde{W}_{u,k}$ are UB with trigger of events for all $k_i \geq k_0 + T$, or alternatively, for all $k \geq k_0 + \bar{T}$ since k_i is a subsequence of k and, hence, $\bar{T}$ is a function of T. Therefore, the Lyapunov function will converge it its ultimate value.

Remark A.1: From both the Cases, the UB for system state, NN weight estimation errors of identifier, critic and actor NNs are found to be $b_x = \max(b_{1,M}^{x}, b_{2,M}^{x})$, $b_{\tilde{W}_I}$, $b_{\tilde{W}_V}$ and $b_{\tilde{W}_u}$, respectively. The bounds b_x, $b_{\tilde{W}_I}$, $b_{\tilde{W}_V}$ and $b_{\tilde{W}_u}$ are function of learning parameters α_I, α_V $, \alpha_u$ and the NN reconstruction error bounds $\varepsilon_{I,M}$, $\varepsilon_{V,M}$, and $\varepsilon_{u,M}$. Hence, a smaller UB for the closed-loop system can be obtained by selecting α_I, α_V, α_u properly and increasing the number of neurons in the NN to reduce $\varepsilon_{I,M}$, $\varepsilon_{V,M}$, and $\varepsilon_{u,M}$.

Finally, to show the convergence of estimated value function and control input to their optimal values, subtract (22) from (20) and (39) from (36) to get

$$\| V^* - \hat{V} \| = \tilde{W}_{V,k}^T \varphi(x_k, k) + \hat{W}_{V,k}^T \left(\varphi(x_k, k) - \varphi(\breve{x}_k, k) \right) + \varepsilon_{V,k}$$

$$\leq b_{\tilde{W}_V} \varphi_M + \hat{W}_{V,\max}^T C_\varphi \sigma_{ET,\max} b_x + \varepsilon_{V,M} \equiv b_V, \tag{A.19}$$

and

$$\| u_k^* - u_k \| = \tilde{W}_{u,k}^T \sigma_u(x_k, k) + \hat{W}_{u,k}^T \left(\sigma_u(x_k, k) - \sigma_u(\breve{x}_k, k) \right) + \varepsilon_{u,k}$$

$$\leq b_{\tilde{W}_u} \sigma_{u,M} + \hat{W}_{u,\max} C_{\sigma_u} \sigma_{ET,\max} b_x + \varepsilon_{u,M} \equiv b_u, \tag{A.20}$$

where $\hat{W}_{V,\max} = \max_k \{\hat{W}_{V,k}\}$ and $\hat{W}_{u,\max} = \max_k \{\hat{W}_{u,k}\}$ are the maximum estimated values for the critic and actor NNs. The maximum value of the event-trigger threshold coefficient is denoted by $\sigma_{ET,\max}$. The constants C_φ and C_{σ_u} are the Lipschitz constants for the critic and actor NN activation functions, respectively. Note that bounds b_V and b_u depend on the UB of the system state vector b_x, NN weights $b_{\tilde{W}_V}$ and $b_{\tilde{W}_u}$, which are small as mentioned in Remark A.1. Therefore, b_V and b_u are small constants and the estimated control input converge to the near optimal value. ■

Proof of Theorem 3: Consider the event-trigger error (2) $e_{ET,k} = x_k - \breve{x}_k$. The error dynamics, $e_{ET,k+1} = x_{k+1} - \breve{x}_{k+1}$, the by using the closed-loop system dynamics (47) is upper bounded by

$$\left\| e_{ET,k+1} \right\| \leq M_i \left\| e_{ET,k} \right\| + N_i, \quad k_i < k < k_{i+1}, i = 1, 2, \cdots \tag{A.21}$$

where $N_i = ((\sqrt{K^*} + 1) \left\| x_{k_i} \right\| + g_M (\sigma_{u,M} \left\| \tilde{W}_{u,k_i} \right\| + \varepsilon_{u,M}))$ and $M_i = (\sqrt{K^*} + g_M C_{\sigma_u} \left\| \hat{W}_{u,k_i} \right\|)$, $i = 1, 2, \cdots$ with $0 < K^* < 1/2$.

Remark A.2: The variables M_i and N_i are piecewise constant functions since $\tilde{W}_{u,k_i}$, $\hat{W}_{u,k_i}$ and x_{k_i} are constant for each i^{th} inter-event time. Hence, the error $e_{ET,k}$ is also a piece wise continuous function.

By comparison lemma [20], the solution of the inequality (A.21) is bounded above as

$$\left\|e_{ET,k}\right\| \le \sum_{j=k_i}^{k-1} M_i^{\,k-j-1} N_i = \left(N_i M_i^{\,k-k_i} - N_i\right)/\left(M_i - 1\right), \tag{A.22}$$

for $k_i < k < k_{i+1}, i = 1, 2, \cdots$. The lower bound on the inter-trigger times for i^{th} inter-event duration, $\delta k_i = k_{i+1} - k_i$, is the time it takes $\left\|e_{ET,k}\right\|$ in (A.22) to reach the minimum threshold, $\sigma_{ET,\min}$ for all $k \in \mathbb{N}$. It is computed using (50) given as

$$\min_{k\in\mathbb{N}}\{\sigma_{ET}\left\|x_k\right\|\} = \left(\sqrt{(1-2K^*)\Gamma_{ET}\big/ 4g_M^2 C_{\sigma_u}^2 \hat{W}_{u,\max}^2}\right) b_x = \sigma_{ET,\min}, \tag{A.23}$$

where b_x is the lower bound of the system state for an event to trigger as in (52). The weight matrix $\hat{W}_{u,\max} = \max_k \left\{\hat{W}_{u,k}\right\}$ is the maximum value of the actor NN weight estimates for all $k \in \mathbb{N}$. The maximum value of the NN weight matrix $\hat{W}_{u,\max}$ exists since the weight estimates are bounded for all time.

The triggering instants are decided by the violation of the event-trigger condition. Thus at k_{i+1} for i^{th} interval, it holds that $\left\|e_{ET,k_{i+1}}\right\| = \sigma_{ET,\min}$. Therefore, from (A.22) we get

$$\left(N_i M_i^{\,k_{i+1}-k_i} - N_i\right)/\left(M_i - 1\right) \ge \sigma_{ET,\min}, i = 1, 2, \cdots. \tag{A.24}$$

Solving the above inequality, the lower bound on the inter-event times found to be

$$\delta k_i \ge \ln\left(1 + \left(1/N_i\right)\left(\left(M_i - 1\right)\sigma_{ET,\min}\right)\right)\Big/\ln\left(M_i\right), i = 1, 2, \cdots \tag{A.25}$$

From (A.25), the minimum value of inter-event time:

$$\delta k_{\min} = \min_{i \in \mathbb{N}} \left(\delta k_i \right) = \min_{i \in \mathbb{N}} \left(\ln(1 + (1/N_i)((M_i - 1)\sigma_{ET,\min})) / \ln(M_i) \right).$$

The inter-event times becomes non trivial, i.e., $\delta k_i > 1$ when

$$\ln(1 + (1/N_i)((M_i - 1)\sigma_{ET,\min})) > \ln(M_i),\ i \in \mathbb{N},$$

is satisfied. ■

IV. NEURAL NETWORK-BASED EVENT-TRIGGERED STATE FEEDBACK CONTROL OF NONLINEAR CONTINUOUS-TIME SYSTEMS

Avimanyu Sahoo, Hao Xu and S. Jagannathan

Abstract — This paper presents a novel approximation based event-triggered control of multi-input multi-output (MIMO) uncertain nonlinear continuous-time systems in affine form. The controller is approximated by using a linearly parameterized neural network (NN) in the context of event-based sampling. After revisiting the NN approximation property in the context of event-based sampling, an event-trigger condition is proposed by using the Lyapunov technique to reduce the network resource utilization and to generate the required number of events for the NN approximation. In addition, a novel weight update law for aperiodic tuning of the NN weights at triggered instants is proposed to relax the knowledge of complete system dynamics and to reduce the computation when compared to the traditional NN-based control. Nonetheless, a non-zero positive lower bound for the inter-event times is guaranteed to avoid accumulation of events or Zeno behaviour. For analysing the stability, the event-triggered system is modelled as a nonlinear impulsive dynamical system and the Lyapunov technique is used to show local ultimate boundedness of all signals. Further, in order to overcome the unnecessary triggered events when the system states are inside the ultimate bound, a dead-zone operator is used to reset the event-trigger error to zero. Finally, the analytical design is substantiated with numerical results.

Keywords- Adaptive control, approximation, event-triggered control, neural network control.

1. INTRODUCTION

Growing interest in the networked control system (NCS) has given rise to an alternate control paradigm known as event-triggered control (ETC) [1]-[18] in order to reduce communication traffic and save computational load on the processors. Instead of transmitting and executing the controller in a traditional periodic sampled manner, the ETC approach provides a mechanism for deciding the sampling instants without compromising the desired performance. The analytically designed trigger condition allows the system error to grow before deciding the transmission instant without affecting the system's stability requirements. This in turn reduces the communication and computation. In recent times, various event-trigger approaches have been presented in the literature [2]-[17], and different formulations have been introduced to analyze system stability. In general, the Lyapunov direct method is utilized to guarantee stability by designing an event-trigger condition.

Among the earlier works, the author in [2] presented an event-triggered control scheme by assuming the input-to-state stability (ISS) of the system with respect to measurement error. Further an event-trigger condition is developed for deciding the trigger instants to execute the controller with a desired closed-loop performance. A lower bound on the inter-event times is also guaranteed to avoid the accumulation point. The traditional ETC [2]-[4] is further extended to a model-based scheme [5]-[8], which reduces the communication network traffic more effectively demanding a higher computation. The ETC also finds its application in large scale and decentralized systems [9]-[11]. An extension to the ETC approach is the self-triggered control design [12]-[15]

where the trigger instants are determined by the past state information and, hence, continuous monitoring of current state is not required.

In these previous works [2]-[7], a known system dynamics have been considered for the ETC design both for linear and nonlinear systems with a few exceptions [8], [17]. In [8], the authors considered known uncertainty for the system and developed a model-based event-triggered control scheme. Further, in [17], an L_1 adaptive control scheme is proposed where the nominal system dynamics are considered known, and uncertainties are compensated for by using an adaptive term tuned with a projection-based tuning law. On the other hand, in our previous preliminary work [18], the complete knowledge of the system dynamics were relaxed by using neural network (NN) based approximation of system dynamics while a zero-order-hold (ZOH) was used for the controller implementation.

In contrast, this paper introduces the development of ETC of MIMO nonlinear continuous time systems in affine form when the system and the controller are separated by an ideal communication network with no delays and packet losses. Instead of approximating the unknown nonlinear functions of the system dynamics by using two NNs [18], the controller is approximated by using a *linearly parameterized* NN in the event-triggered context under the assumption that the system states are measurable. An event-trigger condition based on system state and estimated NN weight is designed to orchestrate the transmission of state vector and control input between the plant and controller. Since the approximation is carried out using the event-based state vector, the event-trigger condition is made adaptive in order to attain a trade-off between resource utilization and function approximation. In addition, the NN weights and the control

inputs are only updated at the trigger instants, which are aperiodic in nature and held until the next update. Consequently, the proposed scheme reduces the overall computation when compared to the traditional NN schemes [19] where weights are updated periodically.

In addition, to analyze the system stability and design the event-trigger condition, the nonlinear impulsive dynamical model of the closed-loop dynamics is considered. The well-developed Lyapunov approach for the nonlinear impulsive dynamical system [20]-[21] is utilized to study inter-event and event time behavior, and used to prove the local ultimate boundedness (UB) of the system state and the NN weight estimation errors. An NN-based control design for a traditional impulsive dynamical system is presented by the authors in [21]. In contrast to [21], in this paper, we modelled the closed-loop event-triggered system as an impulsive dynamical system to analyze stability and performance.

The main contributions of this paper include: (1) the design of an NN-based event-triggered control of uncertain nonlinear continuous-time MIMO systems in affine form, (2) the design of an online approximate controller in the event-triggered context, (3) development of aperiodic event-based NN weight update law to reduce computation, (4) design of novel adaptive event-trigger condition for uncertain nonlinear dynamics to facilitate approximation and to maintain system stability and performance while reducing the transmission, and (5) demonstration of closed-loop stability using the impulsive dynamical system formulation.

The remaining part of the paper is organized as follows. Section 2 discusses the preliminaries. Section 3 presents the state feedback design of the event-triggered control followed by the discussion on non-zero positive lower bound on the inter-event times in

Section 4. The analytical results are verified using numerical example in Section 5 and conclusions are presented in Section 6. The Appendix provides the detailed proofs for the lemmas and theorems.

2. PROBLEM FORMULATION

First the notations used in this paper are briefly introduced followed by the stability notions. Subsequently, a brief background on traditional ETC is presented along with the problem formulation. Finally, the NN based function approximation is revisited in the context of event-based sampling.

2.1 NOTATION

Let $\Re = (-\infty, \infty)$ be the set of real numbers. $\Re^+ = [0, \infty)$ becomes the set of nonnegative real numbers, and $\Re^n$ is the n-dimensional Euclidean space. For a vector, $x \in \Re^n$, we denote $\|x\| \triangleq \sqrt{x^T x}$ its 2-norm. For a matrix, $A \in \Re^{n \times m}$, $\|A\| \triangleq \sqrt{\sum_{i=1}^{n} \sum_{j=1}^{m} |a_{ij}|^2}$ denotes the Frobenius norm. The transpose of A is denoted by $A^T \in \Re^{m \times n}$ and $tr\{\bullet\}$ is the trace operator of a square matrix. Let $W = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ be a matrix, then $\mathbf{vec}(W) = [a \quad b \quad c \quad d]^T$ is the vectorization of the matrix W and $\boldsymbol{vec}(W)^T \boldsymbol{vec}(W) = tr\{W^T W\}$. For a square matrix, $A \in \Re^{n \times n}$, we denote $\lambda_{\max}(A)$ and $\lambda_{\min}(A)$ represent the maximum and minimum eigenvalues of A.

2.2 STABILITY NOTION

Consider a nonlinear impulsive dynamical system represented as

$$\dot{\xi} = F_c(\xi), \ \xi(0) = \xi_0, \ \xi \in \mathscr{C} \subset \mathscr{D}, \ \xi \notin \mathscr{Z}, \tag{1}$$

$$\Delta\xi = F_d(\xi), \ \xi \in \mathscr{D}, \ \xi \in \mathscr{Z}, \tag{2}$$

where $\xi \in \mathcal{D} \subset \Re^{n_\xi}$ is the state vector, $\mathcal{C} \subset \mathcal{D}$ and $\mathcal{Z} \subset \mathcal{D}$, which are, respectively, the flow and the jump sets, and $\mathcal{D}$ is an open set with $0 \in \mathcal{D}$, $\Delta\xi = \xi(t^+) - \xi(t)$ where $\xi(t^+) = \lim_{\varepsilon \to 0} \xi(t+\varepsilon)$. The nonlinear functions $F_c(\xi) \in \Re^{n_\xi}$ and $F_d(\xi) \in \Re^{n_\xi}$ are respectively the continuous and reset dynamics of the impulsive dynamical system. Next, the definitions are stated.

Definition 1[20]: The nonlinear impulsive dynamical system (1) and (2) is locally bounded if there exists a $\gamma > 0$ such that, for every $\delta \in (0, \gamma)$, there exists $\varepsilon = \varepsilon(\delta) > 0$ such that $\|\xi(0)\| < \delta$ implies $\|\xi(t)\| < \varepsilon,\ t \geq 0$.

Definition 2[20]: The nonlinear state dependent impulsive dynamical system (1) and (2) is *locally UB* with bound ε if there exists $\gamma > 0$ such that, for every $\delta \in (0, \gamma)$ there exists $T = T(\delta, \varepsilon) > 0$ such that $\|\xi(0)\| < \delta$ implies $\|\xi(t)\| < \varepsilon$, $t \geq T$ and *globally UB* with bound ε if, for every $\delta \in (0, \infty)$, there exists, $T = T(\delta, \varepsilon) > 0$, such that $\|\xi(0)\| < \delta$ implies $\|\xi(t)\| < \varepsilon,\ t \geq T$.

Definition 3[23]: A continuous function $\alpha : [0, a) \to \Re^+$ is said to belong to class $\mathcal{K}$ if it is strictly increasing and $\alpha(0) = 0$. It is said to belong to class $\mathcal{K}_\infty$ if $a = \infty$ and $\alpha(r) \to \infty$ as $r \to \infty$.

To prove the ultimate boundedness of the impulsive dynamical systems the following Lemma will be used.

Lemma 1 [20]: Consider the impulsive dynamical system (1) and (2). Assume that the jumps occur at distinct time instants and there exists a continuously differentiable function $V : \mathcal{D} \to \Re$ and class $\mathcal{K}$ functions $\alpha(\bullet)$ and $\beta(\bullet)$ such that

$$\alpha\left(\|\xi\|\right) \le V\left(\xi\right) \le \beta\left(\|\xi\|\right),\ \xi \in \mathcal{D}, \tag{3}$$

$$\frac{\partial V\left(\xi\right)}{\partial \xi} F_c\left(\xi\right) < 0,\ \xi \in \mathcal{D},\ \xi \notin \mathcal{Z},\ \|\xi\| > \mu, \tag{4}$$

$$V\left(\xi + F_d\left(\xi\right)\right) - V\left(\xi\right) \le 0,\ \xi \in \mathcal{D},\ \xi \in \mathcal{Z},\ \|\xi\| > \mu, \tag{5}$$

where $\mu > 0$ is such that $\mathcal{B}_{\alpha^{-1}(\beta(\mu))} = \left\{\xi \in \Re^{n_\xi} : \|\xi\| < \alpha^{-1}\left(\beta\left(\mu\right)\right)\right\} \subset \mathcal{D}$ with $\eta > \beta\left(\mu\right)$. Further, assume $\theta \triangleq \sup_{\xi \in \bar{\mathcal{B}}_\mu \cap \mathcal{Z}} V\left(\xi + F_d\left(\xi\right)\right)$ exists. Then the nonlinear state dependent impulsive dynamical system (1) and (2) is UB with bound $\Xi \triangleq \alpha^{-1}\left(\eta\right)$ where $\eta \triangleq \max\left\{\beta\left(\mu\right), \theta\right\}$. Furthermore, $\limsup_{t\to\infty} \|\xi(t)\| \le \alpha^{-1}\left(\beta(\mu)\right)$.

In the next subsection, the problem formulation along with a brief background on the traditional event-triggered control will be introduced.

2.3 BACKGROUND AND PROBLEM FORMUATION

Consider the multi-input multi-output (MIMO) nonlinear uncertain continuous-time system represented in the affine form as

$$\dot{x} = f(x) + g(x)u,\ x(0) = x_0, \tag{6}$$

where $x = [x_1 \quad x_2 \quad \cdots \quad x_n]^T \in \Re^{n_x}$ and $u \in \Re^{m_u}$ are the state and input vectors of the system (6), respectively. The nonlinear vector function, $f(x) \in \Re^{n_x}$, and the matrix function, $g(x) \in \Re^{n_x \times m_u}$, represent the internal dynamics and control coefficient function, respectively. The following assumption on system dynamics is needed in order to proceed.

Assumption 1: The system (6) is controllable and input-to-state linearizable [23]. The internal dynamics, $f(x)$ and control coefficient $g(x)$ are considered unknown with the control coefficient matrix, $g(x)$, bounded above in a compact set for all $x \in \Omega_x \subset \Re^{n_x}$, satisfying $\|g(x)\| \le g_{\max}$ with $g_{\max} > 0$ being a known positive constant [19].

The input-to-state linearizable assumption is satisfied by a wide variety of practical systems such as a robot manipulator, mass damper system and many others. For these classes of controllable nonlinear systems [23] in affine form with complete knowledge of system dynamics, $f(x)$ and $g(x)$, it is demonstrated that there exists an ideal control input u_d for the system (6) of the form

$$u_d = K(x), \tag{7}$$

which renders the closed-loop system asymptotically stable where $K(x)$ is a function of system state vector. The linear closed-loop dynamics can be represented by

$$\dot{x} = Ax, \tag{8}$$

where A is a Hurwitz matrix and can be designed as per the closed-loop performance requirement. By converse Lyapunov theorem [23], an asymptotically stable system admits a Lyapunov function, $V(x): \Omega_x \to \Re$, which satisfies the following inequalities

$$\alpha_1(\|x\|) \le V(x) \le \alpha_2(\|x\|), \tag{9}$$

$$\dot{V}(x) \le -\alpha_3(\|x\|), \tag{10}$$

where α_1, α_2 and α_3 are class $\mathcal{K}$ functions.

Moreover, considering a standard quadratic Lyapunov function, $V(x) = x^T P x$, for the closed-loop system (8), the class $\mathcal{K}$ functions are expressed as $\alpha_1(\|x\|) = \lambda_{\min}(P)\|x\|^2$,

$\alpha_2(\|x\|) = \lambda_{\max}(P)\|x\|^2$ and $\alpha_3(\|x\|) = \lambda_{\min}(Q)\|x\|^2$. The matrices $P \in \Re^{n_x \times n_x}$ and $Q \in \Re^{n_x \times n_x}$ are symmetric, positive definite, and satisfy the Lyapunov equation given by

$$A^T P + PA = -Q. \tag{11}$$

In the case of a traditional NCS, the state vector, x, and the control input, u, are transmitted with a fixed sampling interval T_s. On the other hand, in an event-trigger context, the system state vector is sampled and transmitted at the event-trigger instants only.

Let $\{\tau_k\}$, for $k = 1, 2, ..$ be a monotonically increasing sequence of time instants with $\tau_0 = 0$ such that $\tau_{k+1} > \tau_k$ and $\tau_k \to \infty$ as $k \to \infty$ represent when the events are triggered and the system states, $x(\tau_k)$, and control inputs, $u(\tau_k)$, are transmitted. The event-based/transmitted state and the control input vectors are held, respectively, at the controller and plant by the ZOHs. It is important to note that τ_k is a function of system state x and the last transmitted system state, $\breve{x} = x(\tau_k)$, $\tau_k < t \le \tau_{k+1}$, and is aperiodic in nature.

Define the event-trigger error, $e_s \in \Re^{n_x}$, as

$$e_s = x - \breve{x}, \tau_k < t \le \tau_{k+1}. \tag{12}$$

The trigger instant, τ_k, is determined by the event-trigger condition consists of the event-trigger error (12) and a state dependent threshold. Once the event-trigger error exceeds the threshold (time instant, $t = \tau_k$), the sensor and trigger mechanism initiates the transmission of the current state vector x. The last held event-based state vector $\breve{x}$ jumps to the new value, i.e., $\breve{x}^+ = x,$ for $t = \tau_k$ and held for $\tau_k < t \le \tau_{k+1}$ where $\breve{x}^+ = \breve{x}\left(\tau_k^+\right)$ and

τ_k^+ is the time instant just after τ_k. The event-trigger error is then reset to zero for the next event to occur, i.e.,

$$e_s^+ = 0 \text{ for } t = \tau_k . \tag{13}$$

Since the system dynamics $f(x)$ and $g(x)$ are considered unknown, the implementation of the controller (4) is not possible. Further, in the event-based sampling and transmission context, the intermittent availability of the system state vector at the controller precludes the traditional NN based approximation with a periodic update of the NN weights. Therefore, the NN function approximation property is revisited under the event-based sampling and transmission.

2.4 FUNCTION APPROXIMATION

By the universal approximation property of NN, any continuous function $f(x)$ can be approximated over a compact set for all $x \in \Omega_x \subset \Re^{n_x}$ up to a desired level of accuracy ε_f by the selection of suitable activation functions and an adequate number of hidden layer neurons. Alternatively, there exists an unknown target weight matrix W such that $f(x)$ in a compact set can be written as

$$f(x) = W^T \varphi(V^T x) + \varepsilon_f(x), \tag{14}$$

where $W \in \Re^{l \times b}$ and $V \in \Re^{a \times l}$ represent the target NN weight matrix for the output and input layers, respectively, and defined as

$$(W, V) = \arg \min_{(W,\,V)} [\sup_{x \in \Omega_x} \left\| W^T \varphi(V^T x) - f(x) \right\|]. \tag{15}$$

The activation function $\varphi(\bullet) : \Re^a \to \Re^l$ is a hyperbolic tangent activation function and given by $\varphi(\bullet) = e^{2\bar{x}} - 1 / e^{2\bar{x}} + 1$ with $\bar{x} = V^T x$. The term $\varepsilon_f(x) \in \Re^{n_x}$ is the traditional

reconstruction error and the constants *l*, *a*, and *b* are the number of neurons in the hidden layer, number of input and output of the NN, respectively.

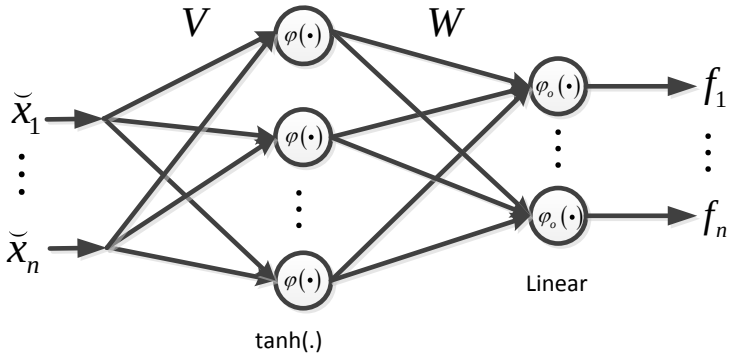

Figure 1. Neural network structure with event-based activation function.

In this paper we will consider the *linearly parameterized* [24] NNs, as shown in Figure 1, for approximating the unknown function as in (14) where the output layer weights $W \in \Re^{l \times b}$ are updated while the input layer weight matrix $V \in \Re^{a \times l}$ is initialized at random and held. This *linearly parameterized* NN is also known as random vector functional link networks (RVFL) [24]. The activation function $\varphi(V^T x)$ forms a basis for the unknown function and the universal approximation property is retained [24]. The output layer activation functions, $\varphi_o(\bullet)$, are selected to be purely linear.

With intermittent event-based transmission of the system state vector, x, the universal NN approximation property can be extended to achieve a desired level of accuracy by properly designing a trigger condition. The trigger condition will generate required number of events for the availability of system state vector for approximation.

The theorem introduced next extends the approximation property of NN for event-based sampling.

Theorem 1: Let $\vec{f}(x):\Omega_x \to \Re^{n_x}$, be smooth and uniformly continuous function in a compact set for all $x \in \Omega_x \subset \Re^{n_x}$. Then, there exists a single layer NN with sufficient number of neurons such that the function $\vec{f}(x)$ can be approximated with constant weights and event-driven activation function, such that

$$\vec{f}(x) = W^T \varphi(\breve{x}) + \varepsilon_e(\breve{x}, e_s), \tag{16}$$

where $W \in \Re^{l \times n_x}$ is the target NN weight matrix with l being the number of hidden-layer neurons, $\varphi(\breve{x})$ is the bounded event-driven activation function, and $\varepsilon_e(\breve{x}, e_s)$ is the event-driven NN reconstruction error with $\breve{x}$ representing the last event sampled state held at the ZOH.

Proof: Refer to Appendix.

Remark 1: From the proof of Theorem 1, the event-based reconstruction error $\varepsilon_e(\breve{x}, e_s) = W^T\left[\varphi(\vartheta(\breve{x}, e_s)) - \varphi(\breve{x})\right] + \varepsilon_{\vec{f}}(\vartheta(\breve{x}, e_s))$ where $\vartheta(\breve{x}, e_s) = \breve{x} + e_s$ is a function of the traditional reconstruction error $\varepsilon_{\vec{f}}(\bullet)$ and event-trigger error e_s as in (12). A small event-based reconstruction error $\varepsilon_e(\breve{x}, e_s)$ can be observed by increasing the frequency of event-based samples. This requires a suitable event-trigger condition for obtaining both approximation accuracy and a reduction in computation. A small event-based reconstruction error means a higher number of events, which results in more computations and transmissions. Hence, a tradeoff exists between reconstruction error and transmission.

3. ADAPTIVE EVENT-TRIGGERED STATE FEEDBACK CONTROL

In this section, a state-feedback design of the NN-based adaptive ETC is proposed.

3.1 STRUCTURE

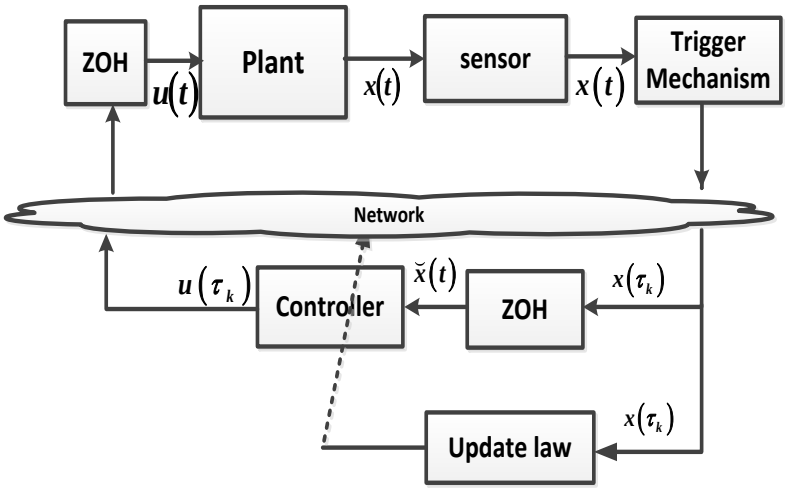

Figure 2. Structure of the adaptive state feedback ETC system.

The structure of the proposed adaptive ETC scheme with a communication network between the plant and the controller is depicted in Figure 2. Further, for simplicity, the following assumption regarding the network is considered.

Assumption 2: The communication network between the plant and the controller is ideal [3], [17], i.e., the networked induced delays including the computational delay and the packet losses are not present.

In the proposed scheme, a smart sensor and trigger mechanism is included at the plant to decide the event-trigger instants by evaluating the event-trigger condition continuously. At the violation of the event-trigger condition, the state vector is

transmitted first and then the controller is updated and transmitted to the plant. The ZOHs are used to hold the last transmitted state and control input, respectively, at the controller and the plant until the next transmission is received.

Since, the system dynamics are considered unknown the control input is approximated by using a NN in an event- sampling context. Further, the NN weights are updated in an aperiodic manner at every trigger instant and held during the inter-event durations. In order to achieve desired approximation accuracy, an adaptive event-trigger condition is designed to generate the required number of events during the learning phase. Thus, the event-trigger condition becomes a function of the NN weight estimates and the system state vector, whereas in the traditional ETC design, it is a function of system state only [2]-[3]. Therefore, to evaluate the event-trigger condition locally, without transmitting the estimated NN weights, the NN weights are updated at both the trigger mechanism and controller in synchronism. This increases the computation but due to the event-based aperiodic update at both places the overall computation reduces. Next, the event-triggered controller design is presented.

3.2 CONTROLLER DESIGN

In this subsection, the approximation-based event-triggered controller design is presented. By the universal approximation property of the NNs, the ideal control input (4) is written as

$$u_d = W_u^T \varphi_u\left(\bar{x}\right) + \varepsilon_u\left(x\right), \tag{17}$$

where $W_u \in \Re^{l_u \times m_u}$ is the output layer unknown ideal NN weight matrix, and $\varphi_u(\bar{x}) \in \Re^{l_u}$ is the tangent hyperbolic activation function with $\bar{x} = V_u^T x$. The function $\varepsilon_u(x) \in \Re^{m_u}$ is the

traditional NN reconstruction error, $V_u \in \Re^{n_x \times l_u}$ is the input layer weight matrix and $l = l_u$, $a = n_x$, and $b = m_u$ are the number of neurons in the hidden layer, number of inputs and outputs of the NN, respectively.

Before presenting the approximation-based controller design, the following standard assumptions are introduced for the NN.

Assumption 3[19]: The target weights, W_u, the activation function, $\varphi_u(\bullet)$, and the reconstruction error $\varepsilon_u(\bullet)$ of the NN are upper bounded in compact set such that $\|W_u\| \le W_{u,\max}$, $\|\varphi_u(\bullet)\| \le \varphi_{u,\max}$ and $\|\varepsilon_u(\bullet)\| \le \varepsilon_{u,\max}$ where $W_{u,\max}, \varphi_{u,\max}$, and $\varepsilon_{u,\max}$ are positive constants.

Assumption 4: The NN activation function, $\varphi_u(\bar{x})$, is considered Lipschitz continuous in a compact set for all $x \in \Omega_x \subset \Re^{n_x}$. Then, for every $x \in \Omega_x \subset \Re^{n_x}$, there exists a constant $L_{\varphi_u} > 0$ such that $\left\|\varphi_u(\bar{x}) - \varphi_u(\breve{\bar{x}})\right\| \le L_{\varphi_u} \|x - \breve{x}\|$ is satisfied.

In the event-triggered context, the actual controller uses the event-based state vector $\breve{x}$ held at the ZOH. Hence, by Theorem 1, the actual event-based control input is represented as

$$u = \hat{W}_u^T \varphi_u(\breve{\bar{x}}),\ \tau_k < t \le \tau_{k+1}, \tag{18}$$

where $\hat{W}_u \in \Re^{l_u \times m_u}$ is the estimated NN weight matrix, $\varphi_u(\breve{\bar{x}}) \in \Re^{l_u}$ is the event-based NN activation function where $\breve{\bar{x}} = V_u^T \breve{x}$ is the scaled input to the NN. Since, the last held state, $\breve{x}$ and the NN weights are updated at the event-trigger instants, $t = \tau_k$, the control input is also updated at the trigger instant, and, then, transmitted to the plant and held by the ZOH until the next update is received.

Further, as proposed, the estimated NN weights, $\hat{W}_u \in \Re^{l_u \times m_u}$, are held during inter-event durations $\tau_k < t \le \tau_{k+1}$ and updated at the trigger instants or referred to as jumps at $t = \tau_k$. Therefore, the NN update law during inter-event durations is defined as

$$\dot{\hat{W}}_u = 0, \text{ for } \tau_k < t \le \tau_{k+1}. \tag{19}$$

Further, at the event-trigger instants, the update law is selected as

$$\hat{W}_u^+ = \hat{W}_u - \frac{\alpha_u}{c + \|e_s\|^2} \varphi_u(\bar{x}) e_s^T L - \kappa \hat{W}_u, \; t = \tau_k, \tag{20}$$

where $\hat{W}_u^+ \in \Re^{l_u \times m_u}$ is the updated NN weight estimate just after the trigger instant with $\alpha_u > 0$ being the NN learning rate, $c > 0$ is a positive constant, $L \in \Re^{n_x \times m_u}$ is a design matrix to match the dimension, and $\kappa > 0$ is a positive constant serving the same role as the sigma-modification [25] in the traditional adaptive control. Note that the update law (20) uses traditional activation $\varphi_u(\bar{x})$ since the system state vector, x, is available for the update at the trigger instant.

Next, define the NN weight estimation error as $\tilde{W}_u = W_u - \hat{W}_u$. The weight estimation error dynamics during the flow, by using (19), can be written as

$$\dot{\tilde{W}}_u = \dot{W}_u - \dot{\hat{W}}_u = 0, \text{ for } \tau_k < t \le \tau_{k+1}, \tag{21}$$

while for the jump instant, $t = \tau_k$, the NN weight estimation error dynamics derived from (20) becomes

$$\tilde{W}_u^+ = W_u - \hat{W}_u^+ = \tilde{W}_u + \alpha_u \chi_s \varphi_u(\bar{x}) e_s^T L + \kappa \hat{W}_u, \text{ for } t = \tau_k, \tag{22}$$

with $\chi_s = 1 / \left(c + \|e_s\|^2\right)$.

In the next subsection, we will formulate the closed-loop dynamics of the adaptive ETC system as a nonlinear impulsive dynamical system to analyze the system behavior for both inter-event and event time instants.

3.3 CLOSED-LOOP SYSTEM IMPULSIVE DYNAMICAL MODEL

As per the proposed scheme, the last transmitted state and the NN weights are updated at the trigger instants only. Hence, the closed-loop event-trigger system behaves as an impulsive dynamical system. Assuming the event instants are distinct, i.e., there exists a non-zero lower bound on the inter-event times, $\delta\tau_k = \tau_{k+1} - \tau_k > 0$, which is proven in Section 4, the closed loop dynamics can be formulated in two steps.

The first step towards the impulsive system modeling is to formulate the flow dynamics. The closed-loop system dynamics during the flow interval for $t \in (\tau_k, \tau_{k+1}]$ can be derived by using both (6) and (18), and represented as

$$\dot{x} = f(x) + g(x)\hat{W}_u^T \varphi_u(\bar{\bar{x}}), t \in (\tau_k, \tau_{k+1}]. \tag{23}$$

Adding and subtracting the ideal control input u_d yields

$$\dot{x} = f(x) + g(x)\left(\hat{W}_u^T \varphi_u(\bar{\bar{x}}) + u_d - u_d\right), t \in (\tau_k, \tau_{k+1}]. \tag{24}$$

Recalling the NN approximation of the ideal controller (18) and the ideal closed-loop dynamics (8), (24) becomes

$$\dot{x} = Ax + g(x)\left(\hat{W}_u^T \varphi_u(\bar{\bar{x}}) - W_u^T \varphi_u(\bar{x}) - \varepsilon_u(x)\right), t \in (\tau_k, \tau_{k+1}]. \tag{25}$$

From the definition, $W_u = \hat{W}_u + \tilde{W}_u$, the closed loop dynamics (25) can be written as

$$\dot{x} = Ax - g(x)\left(\tilde{W}_u^T \varphi_u(\bar{x}) + \varepsilon_u(x)\right) + g(x)\left(\hat{W}_u^T \varphi_u(\bar{\bar{x}}) - \hat{W}_u^T \varphi_u(\bar{x})\right), \; t \in (\tau_k, \tau_{k+1}]. \tag{26}$$

Similarly, the dynamics of the last transmitted state vector, $\breve{x}$, held by the ZOH, during the flow interval becomes

$$\dot{\breve{x}} = 0, t \in (\tau_k, \tau_{k+1}]. \tag{27}$$

Further, the flow dynamics of the NN weight estimation error is given by (21).

In the second and final step, it only remains to formulate the reset dynamics to complete the impulsive modeling of the event-triggered system. This consists of the jumps in the system state, i.e.,

$$x^+ = x, \text{ for } t = \tau_k, \tag{28}$$

the last transmitted state held by the ZOH,

$$\breve{x}^+ = x, \text{ for } t = \tau_k, \tag{29}$$

and NN weight estimation error dynamics (22).

From (28), (29) and (22), the reset dynamics for the system are given by

$$\Delta x = x^+ - x = 0, \text{ for } t = \tau_k, \tag{30}$$

$$\Delta \breve{x} = \breve{x}^+ - \breve{x} = x - \breve{x} = e_s, \text{ for } t = \tau_k, \tag{31}$$

and

$$\Delta \tilde{W}_u = \tilde{W}_u^+ - \tilde{W}_u = \alpha_u \chi_s \varphi_u(\breve{x}) e_s^T L + \kappa \hat{W}_u, \text{ for } t = \tau_k. \tag{32}$$

For formulating the impulsive dynamical system, we consider $\xi_s = \begin{bmatrix} x^T & \breve{x}^T & \mathbf{vec}\left(\tilde{W}_u\right)^T \end{bmatrix} \in \Re^{n_{\xi_s}}$ as the augmented states where $\mathbf{vec}\left(\tilde{W}_u\right) \in \Re^{l_u m_u}$ is the vector form of the NN weight estimation error matrix and $n_{\xi_s} = n_x + n_x + l_u m_u$. Now combine (26), (27) and (21) to obtain the flow dynamics as

$$\dot{\xi}_s = F_c^s(\xi_s), \ \xi_s \in \mathcal{C} \subset \mathcal{D}_s, \ \xi_s \notin \mathcal{Z}_s. \tag{33}$$

Next combine (30), (31) and (32) to get the reset dynamics as

$$\Delta\xi_s = F_d^s\left(\xi_s\right),\ \xi_s \in \mathcal{D}_s,\ \xi_s \in \mathcal{Z}_s, \tag{34}$$

for the impulsive dynamical nonlinear system where the nonlinear functions, $F_c^s\left(\xi_s\right)$ and $F_d^s\left(\xi_s\right)$, are defined as

$$F_c^s\left(\xi_s\right) = \begin{bmatrix} H\left(\xi_s\right) \\ 0 \\ 0 \end{bmatrix} \quad \text{and} \quad F_d^s\left(\xi_s\right) = \begin{bmatrix} 0 \\ e_s \\ \mathbf{vec}\left(\alpha_u \chi_s \varphi_u\left(\bar{x}\right) e_s^T L + \kappa \hat{W}_u\right) \end{bmatrix} \quad \text{with}$$

$H\left(\xi_s\right) = Ax - g\left(x\right)\left(\tilde{W}_u^T \varphi_u\left(\bar{x}\right) + \varepsilon_u(x)\right) + g\left(x\right)\left(\hat{W}_u^T \varphi_u\left(\bar{\bar{x}}\right) - \hat{W}_u^T \varphi_u\left(\bar{x}\right)\right)$ and $\Delta\xi_s = \xi_s^+ - \xi_s$.

The set $\mathcal{D}_s \subset \Re^{n_\xi}$ is an open set with $0 \in \mathcal{D}_s$. The flow set $\mathcal{C}_s \subset \mathcal{D}_s$ is defined as $\mathcal{C}_s = \left\{\xi_s \in \mathcal{D}_s : \|e_s\| \le \sigma_s \|x\|\right\}$, $\mathcal{Z}_s \subset \mathcal{D}_s$ is the jump set and defined as $\mathcal{Z}_s = \left\{\xi_s \in \mathcal{D} : \|e_s\| > \sigma_s \|x\|\right\}$ where $\sigma_s \|x\|$ is the event-trigger threshold to be designed next.

3.4 STABILITY ANALYSIS

In this section the stability results of the closed-loop system are established. Before proceeding, the following lemma for the boundedness of the NN weight estimation error both during the flow and the jump instants is necessary.

Lemma 2 *(Boundedness of the NN weight estimation error):* Consider the nonlinear continuous-time system (6) and the controller (18) expressed as a nonlinear impulsive dynamical system (33) and (34). Let Assumptions 1 through 4 be satisfied while the initial NN weights, $\hat{W}_u\left(0\right)$, are initialized in the compact set Ω_{W_u}. Under the assumption that a non-zero positive lower bound on the inter-event times, $\delta\tau_k = \tau_{k+1} - \tau_k > 0$, $k \in \mathbb{N}$

exists, there exist positive constants $\alpha_u > 0$, $0 < \kappa < 1/2$, $\bar{T}$ and T such that the weight estimation error, $\tilde{W}_u$, is bounded during the flow period and ultimately bounded for all $\tau_k > \bar{T}$ or, alternatively $t > T$ when the NN weights are updated by using (19) and (20).

Proof: Refer to the Appendix.

Next we introduce the event-trigger condition given by

$$D\left(\|e_s\|\right) \le \sigma_s \|x\|, \tag{35}$$

where

$$\sigma_s = \Gamma_s q_{\min} \Big/ 4 g_{\max} L_{\varphi_u} \left\|\hat{W}_u\right\| \|P\|, \tag{36}$$

is the threshold coefficient with $0 < \Gamma_s < 1$ and L_{φ_u} is the Lipschitz constants for the activation functions, $q_{\min}$ is the minimum eigenvalue of Q, P is a symmetric positive definite matrix with P and Q satisfying (11), and $D(\bullet)$ is a dead-zone operator defined as

$$D\left(\|e_s\|\right) = \begin{cases} \|e_s\|, \text{ if } \|x\| > B^x_{s,\max}, \\ 0, \text{ otherwise,} \end{cases} \tag{37}$$

where $B^x_{s,\max}$ is the bound for the system state vector x. The system state vector is transmitted to the controller and the updated control input is transmitted to the plant by the violation of the event-trigger condition (35). To ensure $\|\hat{W}_u\|$ in (36) is non zero while evaluating the trigger condition, the previous non zero value of $\|\hat{W}_u\|$ is used when the estimates become zero.

Next, our main result on the local ultimate boundedness of the closed-loop impulsive dynamical system is introduced by utilizing the adaptive event-trigger condition (35) with the help of the Lyapunov approach [20].

Theorem 2 *(Closed-loop stability):* Consider the nonlinear system (6), the control input (18), NN update laws (19) and (20), expressed as an impulsive dynamical system (33) and (34). Let Assumptions 1 through 4 hold. Assume there exists a non-zero positive lower bound on the inter-event times given by $\delta\tau_k = \tau_{k+1} - \tau_k > 0$, $k \in \mathbb{N}$ and the initial NN weight, $\hat{W}_u(0)$, is initialized in the compact set Ω_{W_u}. Then, the closed-loop system state vector ξ_s for any initial condition $\xi_s(0) \in \mathscr{D}_s \subset \Re^{n_\xi}$ is locally ultimately bounded with a bound $\|\xi_s\| \le \Xi$ provided the events are triggered at the violation of the condition (35). Further, the ultimate bound is given by

$$\Xi = \sqrt{\eta / \lambda_{\min}\left(\bar{P}\right)}, \tag{38}$$

where $\bar{P} = diag\{P, P, I\}$ is a positive definite matrix where I is the identity matrix with appropriate dimension and $\eta = \max\left\{\lambda_{\max}\left(\bar{P}\right)\mu_s^2, \theta\right\}$ with $\theta = \sup_{\xi_s \in \bar{\mathscr{B}}_{\mu_s} \cap \mathscr{L}_s} V_s\left(\xi_s + F_d^s\left(\xi_s\right)\right)$, $\mu_s = \max\left\{B_{s,\max}^x, B_{s,\max}^{\breve{x}}, B_{s,\max}^{\tilde{W}}\right\}$ where $B_{s,\max}^x$, $B_{s,\max}^{\breve{x}}$, and $B_{s,\max}^{\tilde{W}}$ are the bounds for the system state, x, the last transmitted state, $\breve{x}$, and the NN weight estimation error, $\tilde{W}_u$, respectively.

Proof: Refer to the Appendix.

Remark 2: The threshold coefficient σ_s of the event-trigger condition (35) is a function of the norm of NN weight estimates $\left\|\hat{W}_u\right\|$ and, hence, adaptive in nature. Since the

weights are updated only at the trigger instant, $\left\|\hat{W}_u\right\|$ is piecewise constant and jumps at the trigger instant $t=\tau_k$, according to the update law (20). This implies that σ_s is also a piecewise constant function and changes at the trigger instant. This variation in σ_s, implicitly depends on the NN weight estimation error, $\tilde{W}_u$ (more details are in Section 4), which generates the required number of triggers for the NN approximation of the control input during the learning phase. Once the NN weight matrix, $\hat{W}_u$, converges close to the unknown constant target weight matrix, W_u, the weight estimates, $\hat{W}_u$ becomes steady; in turn, σ_s becomes a constant like the traditional event-triggered control with complete knowledge of the system dynamics [2]-[16] .

Remark 3: The dead-zone operator $D(\bullet)$ is used to stop the unnecessary triggering of events due to the NN reconstruction error once the state vector reaches and stays within the UB region. This implies that, for an event to trigger, the following two conditions need to be satisfied:

- The system state vector is outside the bound, i.e. $\|x\| > B_{s,\max}^{x}$, and
- The event-trigger condition (35) is violated, i.e., $\|e_s\| > \sigma_s \|x\|$.

Remark 4: The assumption on the non-zero positive lower bound on inter-event times in Theorem 2 is relaxed by guaranteeing a non-zero positive value in Theorem 3, which is discussed in detail in Section 4. In addition, an explicit formula for analyzing the lower bound on the inter-event times when the system state vector $\|x\| > B_{s,\max}^{x}$ to avoid accumulation point is also derived.

Remark 5: From the proof given in the appendix, the system state vector, x, and NN weight estimation error, $\tilde{W}_u$, remain locally UB for all $\tau_k > \bar{T}$ or alternatively, for all $t > T$ where the time T depends on $\bar{T}$. This implies that the control input and the event-trigger error are also locally ultimately bounded. Consequently, all the closed-loop system parameters remain ultimately bounded for all time $t > T$. The next section will present the lower bound on inter-event times.

4. LOWER BOUND ON INTER-EVENT TIMES

In this section, the existence of non-zero positive lower bound on inter-event times is presented in the following theorem. In addition, an explicit formula for the inter-event times is derived.

Theorem 3: Consider the event-triggered system (6) along with the controller (18) represented as an impulsive dynamical system (33) and (34). Let Assumptions 1 through 4 hold and NN weights, $\hat{W}_u(0)$ is initialized in a compact set Ω_{W_u} and updated using (19) and (20) by the violation of event-trigger condition (35). Then, the lower bound on the inter-event times $\delta\tau_k = \tau_{k+1} - \tau_k$ for all $k \in \mathbb{N}$ implicitly defined by (35) is bounded away from zero and is given by

$$\delta\tau_k \geq \left(1/\|A\|\right)\ln\left(1+\left(\|A\|/\nu_{1,k}\right)\sigma_{s,\min}B_{s,\max}^{x}\right) > 0, \tag{39}$$

where $\sigma_{s,\min}$ is the minimum value of the threshold coefficient over all inter-trigger times. Further, $\nu_{1,k} = g_{\max}\left(\left\|\tilde{W}_{u,k}\right\|\varphi_{u,\max} + \varepsilon_{u,\max}\right) + 2g_{\max}L_{\varphi_u}\varphi_{u,\max}\left\|\hat{W}_{u,k}\right\|$ with $\tilde{W}_{u,k}$ and $\hat{W}_{u,k}$ are the NN weight estimation error and weight estimate for k^{th} flow interval.

Proof: Refer to the Appendix.

Furthermore, it is interesting to study the effect of NN weight estimation error $\tilde{W}_u$ on the inter-event times. The following proposition defines a relation between the lower bound on inter-event times $\delta\tau_k$ and the NN weight estimation error, $\tilde{W}_u$.

Proposition 1: Assume the hypothesis in Theorem 3 holds. Then the lower bound on inter-event times also satisfies

$$\delta\tau_k \geq \left(1/\|A\|\right)\ln\left(1+\left(\|A\|/\nu_{M,k}\right)\sigma_{s,\min}B_{s,\max}^{x}\right), \tag{40}$$

where $v_{M,k} = g_{\max}\left(\varphi_{u,\max}(1+2L_{\varphi_u})\left\|\tilde{W}_{u,k}\right\| + \varepsilon_{u,\max}\right) + 2g_{\max}L_{\varphi_u}\varphi_{u,\max}W_{u,\max}$

Remark 6: It is clear from (40) that the lower bound on inter-event times depends on $v_{M,k}$ which is a function of NN weight estimation error $\tilde{W}_u$. During the initial learning phase of the NN, the term $v_{M,k}$ in (40) might become larger for certain initial value $\hat{W}_u(0)$ and lead to smaller inter-event times closer to zero. A proper initialization of the NN weights, $\hat{W}_u(0)$, close to the target will reduce the weight estimation error, $\tilde{W}_u$, and in turn $v_{M,k}$ in (40). This will keep the inter-event times away from zero and reduce the number of transmissions in the initial phase. In addition, as per Lemma 2, the convergence of the NN weight estimation errors to the bound will further increase the inter-event times leading to less resource utilization as this is verified in the simulation results. Next, the analytical design is evaluated by numerical examples**.**

5. SIMULATION RESULTS

In this section, we validate the theoretical design in the Section 3 and 4 by using numerical examples. Two examples are considered to show the effectiveness of the controller design in terms of saving in communication and computational resources. The first example considers a second order system and is an academic example providing an intuitive idea of the analytical design. In addition, the second example emphasizes the practical application point of view by considering a practical industrial example of a two-link robot manipulator.

5.1 EXAMPLE 1

The following single-input second order nonlinear dynamics [16] was chosen for simulation and given by

$$\begin{aligned} \dot{x}_1 &= x_2, \\ \dot{x}_2 &= -x_1^3 - x_2 + u. \end{aligned} \tag{41}$$

The simulation parameters include the initial state vector as $[5 \ -1]^T$ whereas the closed-loop system matrix is given by $A = [0 \quad 1; -3 \quad -4]$ and the positive definite matrix, $Q = diag\{0.1, 0.1\}$. The learning gain, $\alpha_u = 0.001$, $\kappa = 0.001$ $\Gamma_s = 0.9$, $c = 1$ and $L \in \Re^{2\times 1}$ with elements are all one. The ultimate bound for the system state vector was chosen as 0.001. The tangent hyperbolic activation function, $\tanh\left(V_u^T x\right)$, was used in the NN hidden layer with a randomly initialized fixed input weight, V_u, from the uniform distribution in the interval $[0, 1]$. The Lipschitz constant for the activation function was computed to be $L_{\varphi_u} = \|V_u\| = 4.13$. The number of neurons in the hidden layer was chosen

as $l_u = 15$. The NN weight $\hat{W}_u(0)$ was initialized at random from the uniform distribution in the interval $[0,\ 0.01]$. The sampling time chosen for simulation was 0.001 sec.

Figure 3 (a) illustrates the evolution of the state dependent event error and threshold, and in Figure 3 (b), the cumulative number of events occurred. A total number of events triggered was found to be 645, and the events occurred frequently during the initial NN learning phase. This is due to large initial NN weight estimation error, $\tilde{W}_u$ as discussed in Remark 6. Alternatively, the event-trigger condition generates the required number of triggers for the NN to approximate the control input. A proper selection of the initial weights, $\hat{W}_u(0)$, will further reduce the number of initial triggers.

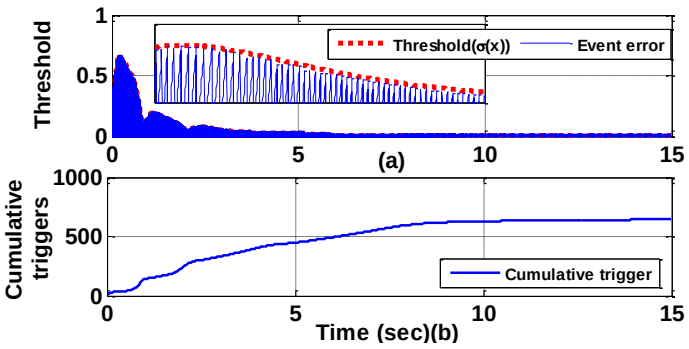

Figure 3. (a) Evolution of the event-trigger threshold; and (b) cumulative number of events

Furthermore, the lower bound on the inter-event times is observed to be 0.002 sec, as shown in Figure 4, implying the existence of a non-zero lower bound on the inter-event times to avoid accumulation point. It is clear from Figure 4 that the inter-event times are gradually increasing along with the convergence of the weight estimation error,

$\tilde{W}_u$, to its ultimate bound, as presented in Proposition 1 and discussed in Remark 6. This elongated inter-event times and reduces resource utilization which is one of the primary objectives of the design.

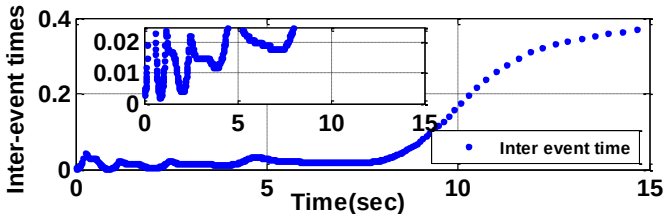

Figure 4. Existence of a nonzero positive lower bound on inter-event times and gradual increase with convergence of NN weight estimates to target.

Figures 5 (a) and (b) depict the convergence of the closed-loop ETC system state vector, and approximated control input. This implies the event-based control input with reduced computation is able to regulate the system state close to zero. Figure 6 shows the convergence of the estimated NN weights with aperiodic weight update. Next, we consider the benchmark example of an MIMO system to evaluate the design.

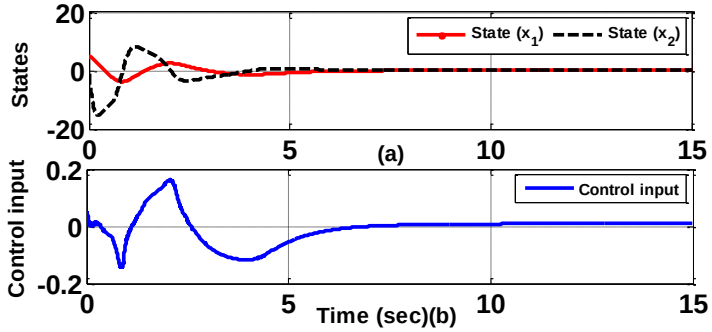

Figure 5. Convergence of (a) system states; and (b) approximated control input.

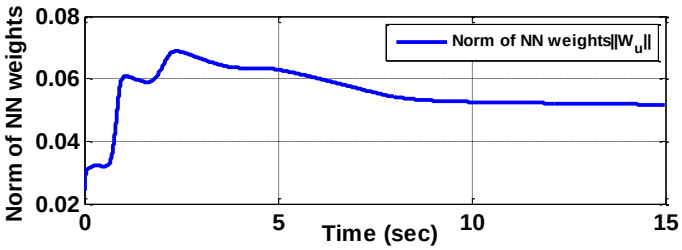

Figure 6. Convergence of the NN weight estimates.

5.2 EXAMPLE 2

A two link robot manipulator is considered whose dynamics are given by

$$\dot{x} = f(x) + g(x)u, \tag{42}$$

where $f(x) = \begin{bmatrix} x_3 \\ x_4 \\ \dfrac{\begin{pmatrix} -(2x_3x_4 + x_4^2 - x_3^2 - x_3^2\cos x_2)\sin x_2 \\ +20\cos x_1 - 10\cos(x_1 + x_2)\cos x_2 \end{pmatrix}}{\cos^2 x_2 - 2} \\ \dfrac{\begin{pmatrix} (2x_3x_4 + x_4^2 + 2x_3x_4\cos x_2 + x_4^2\cos x_2 + 3x_3^2 \\ +2x_3^2\cos x_2 + 20(\cos(x_1 + x_2) - \cos x_1)\times \\ (1+\cos x_2) - 10\cos x_2\cos(x_1 + x_2) \end{pmatrix}}{\cos^2 x_2 - 2} \end{bmatrix}$ and $g(x) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ \dfrac{1}{2-\cos^2 x_2} & \dfrac{-1-\cos x_2}{2-\cos^2 x_2} \\ \dfrac{-1-\cos x_2}{2-\cos^2 x_2} & \dfrac{3+2\cos x_2}{2-\cos^2 x_2} \end{bmatrix}$.

The following simulation parameters were selected for the simulation. The initial state vector is given by $x = \begin{bmatrix} \pi/3 & -\pi/10 & 0 & 0 \end{bmatrix}^T$ while the closed-loop matrix $A = \text{diag}\{-3,\ -4,\ -6,\ -8\}$ and the positive definite matrix was chosen as $Q = \text{diag}\{0.1,\ 0.1,\ 0.1,\ 0.1\}$. The learning gain was selected as $\alpha_u = 0.5$, $\Gamma_s = 0.9$, $\kappa = 0.0015$, $L \in \Re^{4\times 2}$ with elements all one, $g_{\max} = 3$ and $c = 1$. The bound for system state vector was chosen as 0.001. The tangent hyperbolic activation function was used in the hidden layer of the NN with a randomly initialized fixed input weight V_u from the

uniform distribution in the interval $[0,\ 1]$. The Lipschitz constant for the activation function was computed to be $L_{\varphi_u} = \|V_u\| = 3.42$. The number of neurons in the hidden layer was selected as $l = 15$. The NN weight $\hat{W}_u(0)$ was initialized at random from the uniform distribution in the interval $[0,\ 0.01]$. The sampling time chosen for simulation was 0.001 sec.

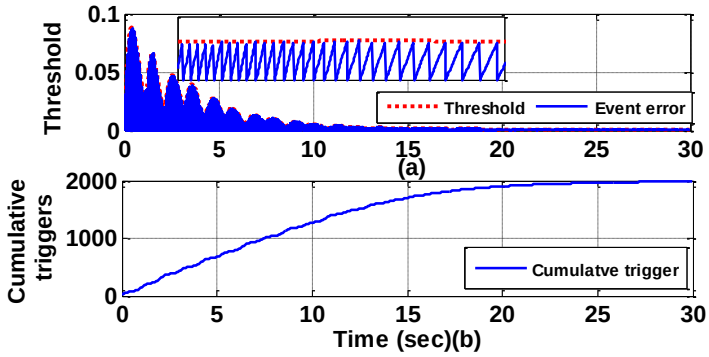

Figure 7. (a) Evolution of the event-trigger threshold; (b) cumulative number of events.

The event-trigger threshold is shown in Figure 7 (a) along with the event-trigger error. The cumulative number of triggered events is illustrated in Figure 7(b), which shows the state vector is only transmitted 2000 times indicating the reduction in communication bandwidth usage when compared to a continuous transmission. Further, the lower bound on the inter-event times is found to be 0.002 sec proven in Theorem 3. In addition, as per Proposition 1, the inter-event times increase with convergence of the NN weight estimates to target as shown in Figure 8.

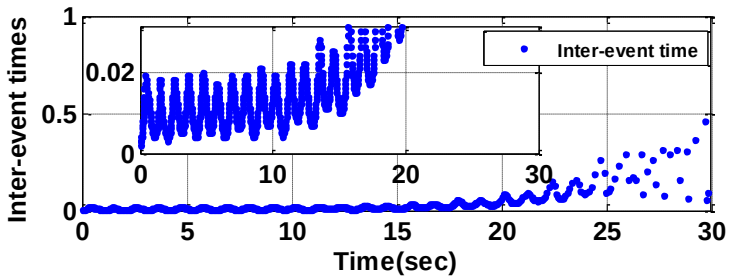

Figure 8. Existence of a nonzero positive lower bound on inter-event times and gradual increase with convergence of NN weight estimates to target.

Further, from Theorem 3 the cumulative number of events depends upon the initial NN weights. The histogram in Figure 9 shows the plot between the norm of initial weights and cumulative number of events. It is clear that the cumulative number of events varies with weight initialization.

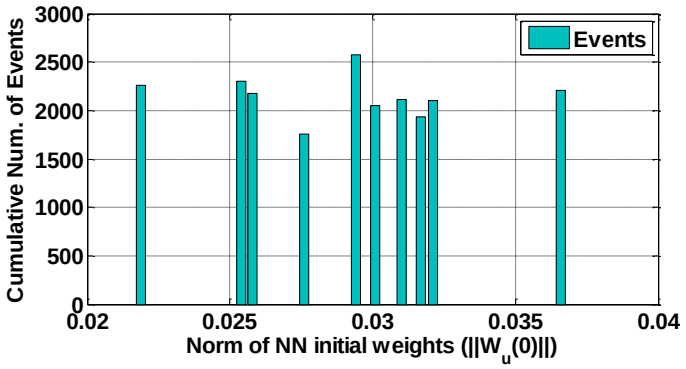

Figure 9. Cumulative number of events with different NN initial weights.

Convergence of the system state and control input is shown in Figures 10 (a) and (b), respectively, implying the event-based controller-regulated system states close to zero. Further, the convergence of the estimated NN weights to target value with aperiodic event-based update law is shown in Figure 11.

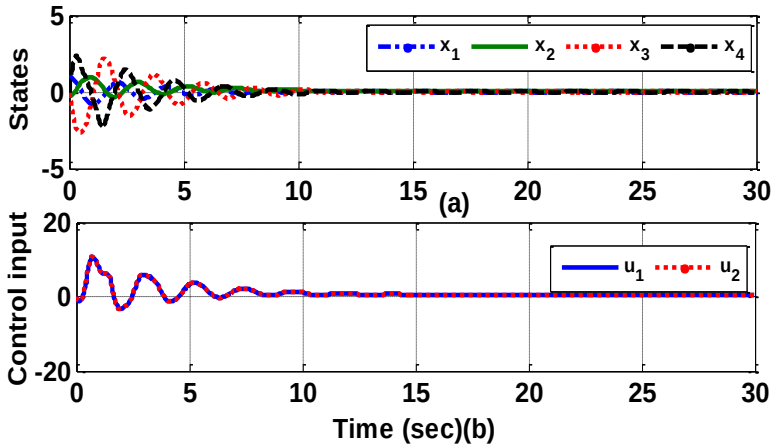

Figure 10. Convergence of (a) system state vectors; and (b) control input.

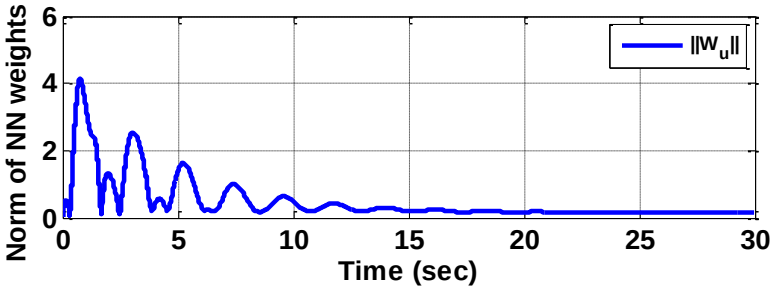

Figure 11. Convergence of the NN weight estimates.

Table 1. Comparison of computational load between traditional periodic and event sampled system.

System		Traditional periodic sampled data system	Event-based non-periodic sampled
Sampling instants		30,000	2000
Number of additions and Multiplications at every sampling instant	NN update law at the controller	10	10
	Controller	3	3
	Update law at the trigger mechanism	0	10
	Trig. Condition (periodic)	0	6
Total number of Computation		390,000	226000

Finally, comparison results in terms of computation and the network traffic between a sampled-data system with a fixed periodic sampling and the event-based sampling is presented in Table 1 and Figure 12 respectively.

Table 1 gives the number of computations observed in terms of addition and multiplications that is needed for realizing both the methods. It is evident that with the event-based system, a 48% reduction in computation when compared to the sample data approach is observed. Further, considering each packetized transmission is of 8 bit data through the ideal network, Figure 12 shows a comparion between the data rate in both the cases. It is clear that the data rate in the case of event-based sampling is lower implying that the needed network bandwidth is less. This verifies the resourcefullness of the event-triggered control design.

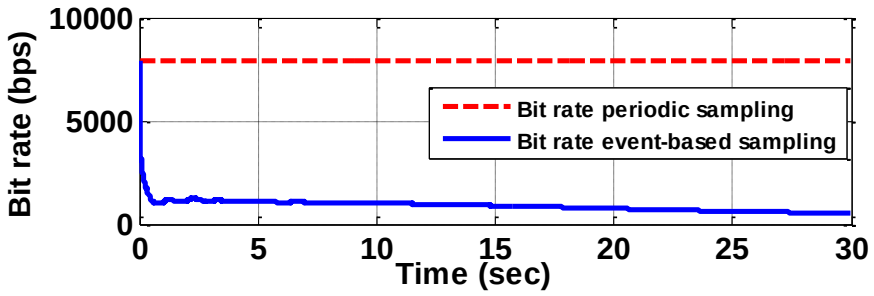

Figure 12. Comparison of data transfer rate between periodic sampled-data and the event-sampled controller for a physical system with a network.

6. CONCLUSIONS

This paper presented an event-triggered stabilization of MIMO uncertain nonlinear continuous-time systems. The control input was directly approximated by using an NN in the context of event-based transmission. Novel event-trigger condition was developed based on the system state vector and NN weight estimate to ensure the reduction in transmission of feedback control signal. The weights were updated in a non-periodic manner at the trigger instants. The controller design guaranteed the desired performance while relaxing the need for system dynamics. Lyapunov analysis confirmed the closed-loop stability. Simulation results confirmed the validity of the control design and reduction in resource utilization.

7. REFERENCES

[1] K. J. Åström and B. M. Bernhardsson, "Comparison of periodic and event based sampling for first-order stochastic systems," in *Proceedings of the 14th IFAC World Congress*, Beijing, China, Jul. 1999, pp. 301–306.

[2] P. Tabuada, "Event-triggered real-time scheduling of stabilizing control tasks," *IEEE Trans. on Automatic Control*, vol. 52, pp. 1680-1685, Sep. 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 Automatic Control*, vol. 57, pp. 1362–1376, Jun. 2012.

[4] R. Postoyan, A. Anta, D. Nesic, and P. Tabuada, "A unifying Lyapunov-based framework for event-triggered control of nonlinear systems," in *Proceedings of the 50th IEEE Conference on Decision and Control and European Control Conference*, Orlando, FL, USA, Dec. 2011, pp. 2559–2564.

[5] W. P. M. H. Heemels and M. C. F. Donkers, "Model-based periodic event-triggered control for linear systems," *Automatica*, vol. 49, no. 3, pp. 698-711, 2013.

[6] J. Lunze and D. Lehmann, "A state-feedback approach to event-based control," *Automatica*, vol. 46, no. 1, pp. 211–215, 2010.

[7] C. Stocker and J. Lunze, "Event-based control of nonlinear systems: An input-output linearization approach," in *Proceedings of the 50th IEEE Conference on Decision and Control and European Control Conference*, Orlando, FL, USA, Dec. 2011, pp. 2541–2546.

[8] E. Garcia and P. J. Antsaklis, "Model-based event-triggered control for systems with quantization and time-varying network delays," *IEEE Trans. on Automatic Control*, vol. 58, pp. 422–434, Feb. 2013.

[9] M. Mazo Jr. and P. Tabuada, "Decentralized event-triggered control over wireless sensor/actuator networks," *IEEE Trans. on Automatic Control*, vol. 56, pp. 2456–2461, Oct. 2011.

[10] X. Wang and M. Lemmon, "Decentralized event-triggered broadcasts over networked control systems," in *Hybrid Systems: Computation and control*, M. Egerstedt and B. Mishra, Eds. Berlin Heidelberg: Springer, 2008.

[11] X. Wang and M. Lemmon, "Event-triggering in distributed networked control systems," *IEEE Trans. on Automatic Control*, vol. 56, pp. 586–601, Mar. 2011.

[12] M. Mazo Jr. and P. Tabuada, "On event-triggered and self-triggered control over sensor/actuator networks," in *Proceedings of the 47th IEEE Conference on Decision and Control*, Cancun, Mexico, Dec. 2008, pp. 435–440.

[13] A. Anta and P. Tabuada, "Self-triggered stabilization of homogeneous control systems," *in Proceedings of the American Control Conference*, Seattle, WA, USA, Jun. 2008, pp. 4129–4134.

[14] A. Anta and P. Tabuada, "To sample or not to sample: Self-triggered control for nonlinear systems," *IEEE Trans. Automatic Control*, vol. 55, pp. 2030–2042, Sep. 2010.

[15] X. Wang and M. Lemmon, "Self-triggered feedback control systems with finite-gain L2 stability," *IEEE Trans. Automatic Control*, vol. 45, pp. 452–467, Mar. 2009.

[16] P. Tallapragada and N. Chopra, "On event triggered tracking for nonlinear systems," *IEEE Trans. on Automatic Control*, vol. 58, pp. 2343–2348, Sep. 2013.

[17] X. Wang and N. Hovakimyan, "L1 adaptive control of event-triggered networked systems," in *Proceedings of the American Control Conference*, Baltimore, MD, USA, Jun. 2010, pp. 2458-2463.

[18] A. Sahoo, H. Xu, and S. Jagannathan, "Neural network-based adaptive event-triggered control of nonlinear continuous-time systems," in *Proceedings of the IEEE International Symposium on Intelligent Control Part of IEEE Multi-Conference on Systems and Control*, Hyderabad, India, Aug. 2013, pp. 35-40.

[19] F. Lewis, S. Jagannathan, and A. Yesildrek, *Neural Network Control of Robot Manipulators and Nonlinear Systems*. London, UK: Taylor and Francis. 1999.

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

[21] T. Hayakawa and W. Haddad, "Stable neural hybrid adaptive control for nonlinear uncertain impulsive dynamical systems," in *Proceedings of the 44th IEEE Conference on Decision and Control, and the European Control Conference*, Seville, Spain, Dec. 2005, pp. 5510–5515.

[22] R. Goebel, R. Sanfelice, and A. Teel, "Hybrid dynamical systems," *IEEE Control System Magazine*, vol. 29, pp. 28–93, Apr. 2009.

[23] H. K. Khalil, *Nonlinear Systems*. Upper Saddle River, NJ, USA: Prentice Hall, 2002.

[24] S. Jagannathan, *Neural Network Control of Nonlinear Discrete-Time System.* Boca Raton, FL, USA: CRC Press, 2006.

[25] P. Ioannou and B. Fidan, *Adaptive Control Tutorial.* Philadelphia, PA, USA: SIAM, 2006.

[26] G. Cybenko, "Approximation by super positions of sigmoidal function," *Mathematics of Control, Signals, and Systems,* vol. 2, no. 4, pp. 304-314, 1989.

[27] T. Dierks and S. Jagannathan, "Optimal control of affine nonlinear continuous-time systems using an online Hamilton-Jacobi-Isaacs formulation," in *Proceedings of the 49th IEEE Conference on Decision and Control*, Atlanta, GA, Dec. 2010, pp. 3048-3053.

APPENDIX

Proof of Theorem 1: Recall universal approximation theorem [26][19], the smooth and uniformly continuous function $\vec{f}(x)$ can be approximated by utilizing a NN with constant weights and time-driven periodic activation function as

$$\vec{f}(x) = W^T \varphi(x) + \varepsilon_{\vec{f}}(x) \ . \tag{A.1}$$

Moreover, the equation (A.1) can be expressed using event-driven activation function (i.e., $\varphi(\breve{x})$) with the event-based state $\breve{x} = x(\tau_k)$ for $\tau_k < t \le \tau_{k+1}$ as

$$\begin{aligned} \vec{f}(x) &= W^T \varphi(x) - W^T \varphi(\breve{x}) + W^T \varphi(\breve{x}) + \varepsilon_{\vec{f}}(x) \\ &= W^T \varphi(\breve{x}) + W^T [\varphi(x) - \varphi(\breve{x})] + \varepsilon_{\vec{f}}(x). \end{aligned} \tag{A.2}$$

Further, from the event-trigger error (12), the current system state x can be represented by using last event-based state and the event-trigger error given by $x = \breve{x} + e_s = \vartheta(\breve{x}, e_s)$. Therefore, the (A.2) can be written as

$$\vec{f}(x) = W^T \varphi(\breve{x}) + \varepsilon_e(\breve{x}, e_s), \tag{A.3}$$

where $\varepsilon_e(\breve{x}, e_s) = W^T [\varphi(\vartheta(\breve{x}, e_s)) - \varphi(\breve{x})] + \varepsilon_{\vec{f}}(\vartheta(\breve{x}, e_s))$. ■

Proof of Lemma 2*:* By the hypothesis of Lemma 2, the events are occurring at discrete time instants, i.e., a nonzero positive lower bound on the inter-event times, exists. Therefore, the proof is carried out, considering the flow and the jump dynamics of the NN weight update law as in Lemma 1, in two different cases as follows.

Case I: (*Boundedness of the weight estimation error during the flow for* $t \in (\tau_k, \tau_{k+1}]$).

This proof is trivial and can be seen by selecting a Lyapunov function $V_s(\tilde{W}_u) = tr\{\tilde{W}_u^T \tilde{W}_u\}$ and the first derivative along the weight estimation error

$$\dot{V}_s(\tilde{W}_u) = 0\,. \tag{A.4}$$

It is clear from (A.4) that weight estimation error remains constant during the flow intervals for $t \in (\tau_k, \tau_{k+1}]$. As the initial weights are finite and the target NN weights are bounded, the NN weight estimation error $\tilde{W}_u$ is constant and bounded during the flow interval. It only remains to prove that the weight estimation error during the jump is also remains bounded and converge to the ultimate bound.

Case II: (*Boundedness of the weight estimation error during jump for* $t = \tau_k$*)*

Consider the same Lyapunov function, as in Case I, for the jump instants

$$V_s(\tilde{W}_u) = tr\{\tilde{W}_u^T \tilde{W}_u\}\,. \tag{A.5}$$

The first difference $\Delta V_s(\tilde{W}_u)$ is written as

$$\Delta V_s(\tilde{W}_u) = tr\{\tilde{W}_u^{T+} \tilde{W}_u^{+}\} - tr\{\tilde{W}_u^T \tilde{W}_u\}\,. \tag{A.6}$$

Along the weight estimation error dynamics (22), the first difference $\Delta V_s(\tilde{W}_u)$ in (A.6) is expressed as

$$\Delta V_s(\tilde{W}_u) = tr\left\{(\tilde{W}_u + \alpha_u \chi_s \varphi_u(\bar{x}) e_s^T L + \kappa \hat{W}_u)^T\ (\tilde{W}_u + \alpha_u \chi_s \varphi_u(\bar{x}) e_s^T L + \kappa \hat{W}_u)\right\} - tr\{\tilde{W}_u^T \tilde{W}_u\}$$

$$= 2\alpha_u \chi_s tr\left\{\tilde{W}_u^T \varphi_u(\bar{x}) e_s^T L\right\} + 2\kappa tr\{\tilde{W}_u^T \hat{W}_u\} + \alpha_u^2 \chi_s^2 tr\left\{\left(\varphi_u(\bar{x}) e_s^T L\right)^T \left(\varphi_u(\bar{x}) e_s^T L\right)\right\} + 2\alpha_u \kappa \chi_s$$
$$\times tr\left\{\left(\varphi_u(\bar{x}) e_s^T L\right)^T \hat{W}_u\right\} + \kappa^2 tr\left\{\hat{W}_u^T \hat{W}_u\right\}.$$

By replacing $\hat{W}_u = W_u - \tilde{W}_u$ from the definition, the first difference leads to

$$\Delta V_s\left(\tilde{W}_u\right) = 2\alpha_u \chi_s tr\left\{\tilde{W}_u^T \varphi_u(\bar{x}) e_s^T L\right\} + 2\kappa tr\left\{\tilde{W}_u^T W_u\right\} - 2\kappa tr\left\{\tilde{W}_u^T \tilde{W}_u\right\}$$
$$+\alpha_u^2 \chi_s^2 tr\left\{\left(\varphi_u(\bar{x}) e_s^T L\right)^T \left(\varphi_u(\bar{x}) e_s^T L\right)\right\} + 2\alpha_u \kappa \chi_s tr\left\{\left(\varphi_u(\bar{x}) e_s^T L\right)^T W_u\right\}$$
$$-2\alpha_u \kappa \chi_s tr\left\{\left(\varphi_u(\bar{x}) e_s^T L\right)^T \tilde{W}_u\right\} + \kappa^2 tr\left\{(W_u - \tilde{W}_u)^T (W_u - \tilde{W}_u)\right\}.$$

By using the bounds from Assumption 3, applying Cauchy-Schwartz (C-S) inequality and applying Young's inequality, and with the fact that $0 \le \chi_s \|e_s\| < 1$ one can arrive at

$$\Delta V_s(\tilde{W}_u) \le 2\alpha_u \varphi_{u,\max}(1+\kappa)\|L\|\|\tilde{W}_u\| - (\kappa - 2\kappa^2)\|\tilde{W}_u\|^2 + B^{W_u}, \tag{A.7}$$

where $B^{\tilde{W}_u} = 2\alpha_u \kappa \varphi_{u,\max}\|L\|W_{u,\max} + (2\kappa^2 + \kappa)W_{u,\max}^2 + \alpha_u^2 \varphi_{u,\max}^2 \|L\|^2$. Defining $a_1 = (\kappa - 2\kappa^2) > 0$ with $0 < \kappa < 1/2$, and $a_2 = 2\alpha_u \varphi_{u,\max}\|L\|(1+\kappa)$, the first difference becomes

$$\Delta V_s\left(\tilde{W}_u\right) \le -a_1\|\tilde{W}_u\|^2 + a_2\|\tilde{W}_u\| + B^{\tilde{W}_u}.$$

Completing the square for $\|\tilde{W}_u\|$ reveals that

$$\Delta V_s\left(\tilde{W}_u\right) \le -(a_1/2)\|\tilde{W}_u\|^2 - \left(\sqrt{a_1/2}\|\tilde{W}_u\| - \left(a_2/\sqrt{2a_1}\right)\right)^2 + B_1^{\tilde{W}_u}$$

where $B_1^{\tilde{W}_u} = \left(a_2^2/2a_1\right) + B^{\tilde{W}_u}$. Since the second term is always negative, it also holds that

$$\Delta V_s\left(\tilde{W}_u\right) \le -(a_1/2)\|\tilde{W}_u\|^2 + B_1^{\tilde{W}_u}. \tag{A.8}$$

From (A.8) the first difference $\Delta V_s(\tilde{W}_u) < 0$ as long as $\|\tilde{W}_u\|^2 > (2B_1^{\tilde{W}_u}/a_1) = B_{\max}^{\tilde{W}_u}$. Hence, by Lyapunov theorem, the NN weight estimation error is ultimately bounded [20] with the trigger of events and for all $\tau_k > \bar{T}$ or alternatively for all $t > T$ where T is function of $\bar{T}$.

Consequently, from Case I and II, the NN weight estimation error remains constant and bounded during the flow period and converges to the ultimate bound with the trigger of events for all $\tau_k > \bar{T}$. Therefore we can conclude that the NN weight estimation error is ultimately bounded for all $t > T$. ■

Proof of Theorem 2: To prove this theorem we will use the Lemma 1. In other words, we need to show that (4) for flow and (5) for reset dynamics hold. For the flow duration, consider a Lyapunov function candidate for $t \in (\tau_k, \tau_{k+1}]$

$$V_u(\xi_u) = \xi_u^T \bar{P} \xi_u, \tag{A.9}$$

with $\bar{P} = diag\{P, P, I\}$ where P is a symmetric positive definite matrix, I is the identity matrix of appropriate dimension. The Lyapunov function (A.9) can be expanded as

$$V_s(\xi_s) = V_s^x + V_s^{\breve{x}} + V_s^{\tilde{W}_u}, \tag{A.10}$$

with $V_s^x = x^T P x$, $V_s^{\breve{x}} = \breve{x}^T P \breve{x}$ and $V_s^{\tilde{W}_u} = \mathbf{vec}(\tilde{W}_u)^T \mathbf{vec}(\tilde{W}_u) = tr\{\tilde{W}_u^T \tilde{W}_u\}$.

The first derivative of the first term in (A.10), $\dot{V}_s^x = \dot{x}^T P x + x^T P \dot{x}$ along the system trajectories (33) can be expressed as

$$\begin{aligned}\dot{V}_s^x &= \left(Ax - g(x)\left(\tilde{W}_u^T \varphi_u(\bar{x}) + \varepsilon_u(x)\right) + g(x)\left(\hat{W}_u^T \varphi_u(\bar{\bar{x}}) - \hat{W}_u^T \varphi_u(\bar{x})\right)\right)^T Px \\ &+ x^T P\left(Ax - g(x)\left(\tilde{W}_u^T \varphi_u(\bar{x}) + \varepsilon_u(x)\right) + g(x)\left(\hat{W}_u^T \varphi_u(\bar{\bar{x}}) - \hat{W}_u^T \varphi_u(\bar{x})\right)\right)\end{aligned}$$

$$= -x^T Q x - 2g(x)\left(\tilde{W}_u^T \varphi_u(\bar{x}) + \varepsilon_u(x)\right)^T Px + 2g(x)\left(\hat{W}_u^T \varphi_u(\bar{\bar{x}}) - \hat{W}_u^T \varphi_u(\bar{x})\right)^T Px,$$

where Q is a positive definite matrix satisfying the Lyapunov equation $A^T P + PA = -Q$. Using Frobenius norm and applying C-S inequality, it also holds that

$$\begin{aligned}\dot{V}_s^x &\le -q_{\min}\|x\|^2 + 2g_{\max}\left(\left\|\tilde{W}_u\right\|\varphi_{u,\max} + \varepsilon_{u,\max}\right)\|P\|\|x\| \\ &+ 2g_{\max}\left\|\left(\hat{W}_u^T \varphi_u(\bar{\bar{x}}) - \hat{W}_u^T \varphi_u(\bar{x})\right)\right\|\|P\|\|x\|,\end{aligned} \tag{A.11}$$

where $q_{\min} > 0$ is the minimum eigenvalue value Q. Again using the Lipschitz continuity from Assumption 4 and separating the cross terms using Young's inequality, $2ab \le (1/p)a^2 + pb^2$, in (A.11), we arrive at

$$\dot{V}_s^x \le -q_{\min}\|x\|^2 + (q_{\min}/2)\|x\|^2 + 2g_{\max}L_{\varphi_u}\left\|\hat{W}_u\right\|\|P\|\|x\|\|e_s\| + B_s^x \tag{A.12}$$

where $B_s^x = (2/q_{\min})\, g_{\max}^2 \|P\|^2 \left(\left\|\tilde{W}_u\right\|\varphi_{u,\max} + \varepsilon_{u,\max}\right)^2$. Note that B_s^x is constant during flow, $t \in (\tau_k, \tau_{k+1}]$, for a fixed k as the NN weight estimation error $\tilde{W}_u$ remains constant.

Recalling the event-trigger condition (35), the first difference in (A.12) becomes

$$\dot{V}_s^x \le -(q_{\min}/2)\|x\|^2 + 2\sigma_s g_{\max}L_{\varphi_u}\left\|\hat{W}_u\right\|\|P\|\|x\|^2 + B_s^x .$$

Substituting σ_s from (36), the first difference leads to

$$\dot{V}_s^x \le -(1-\Gamma_s)(q_{\min}/2)\|x\|^2 + B_s^x . \tag{A.13}$$

Similarly, consider the second term in (A.10), the first derivative, using (33), is given by

$$\dot{V}_s^{\breve{x}} = \dot{\breve{x}}^T P\breve{x} + \breve{x}^T P\dot{\breve{x}} = 0 . \tag{A.14}$$

As the NN weights are not updated the first derivative of the third term, $V_s^{\tilde{W}_u}$, in the Lyapunov function (A.10) also becomes zero as given in (A.4), i.e., $\dot{V}_s^{\tilde{W}_u} = 0$.

Combining (A.13), (A.14) and (A.4), first derivative of the Lyapunov function (A.10) for the overall system during flow for $t \in (\tau_k, \tau_{k+1}]$ becomes

$$\dot{V}_s(\xi_s) \le -\frac{q_{\min}}{2}(1-\Gamma_s)\|x\|^2 + B_s^x \text{ for } t \in (\tau_k, \tau_{k+1}]. \tag{A.15}$$

From (A.15) it is clear that the first derivative $\dot{V}_s(\xi) < 0$ as long as $\|x\| > \sqrt{2B_s^x / q_{\min}(1-\Gamma_s)} \equiv B_1^x$. This implies that the system state is bounded during the flow. Further, since the NN weight estimation error, $\tilde{W}_u$, and the last held state, $\breve{x}$, are constant due to no update, $\tilde{W}_u$ and $\breve{x}$ also remain bounded during the flow.

Note that B_s^x in (A.15) is a function of $\tilde{W}_u$. And, from Lemma 2, the NN weight estimation error $\tilde{W}_u$ remains constant during the flow and converges to the bound $B_{\max}^{\tilde{W}_u}$ for all trigger instants $\tau_k > \bar{T}$. This implies B_s^x will converge to $B_{s,b}^x$ for all $\tau_k > \bar{T}$ or, alternatively, for all $t > T$ where $B_{s,b}^x = \left(2/q_{\min}\right) g_{\max}^2 \left\|P\right\|^2 (B_{\max}^{\tilde{W}_u} \varphi_{u,\max} + \varepsilon_{u,\max})^2$. It follows that, the bound B_1^x for the system states, x, is will converge to $B_{s,\max}^x$ with trigger events for all time $t > T$ where $B_{s,\max}^x = \sqrt{2B_{s,b}^x / q_{\min}\left(1-\Gamma_s\right)}$. It only remains to show that (5) holds for the reset dynamics.

Remark A.1: From (A.15) it is clear that the system state x will remain bounded during the flow for $t \in (\tau_k, \tau_{k+1}]$. As per the reset dynamics (34) of the impulsive dynamical system, $x^+ = x$. Hence, with a finite initial value, the system state vector also remains bounded at the jump instant at $t = \tau_k$ for all $k = 1, 2, 3....$ Further, since $\breve{x}^+ = x$ for $t = \tau_k$, $\breve{x}$ also remains bounded during the jump.

To show that (5) holds during the jump, we select the same Lyapunov function candidate as in (A.9)

$$V_s\left(\xi_s\right) = \xi_s^T \bar{P} \xi_s . \tag{A.16}$$

The first difference in an expanded form is given by

$$\Delta V_s\left(\xi_s\right) = V_s\left(\xi_s^+\right) - V\left(\xi_s\right) = \xi_s^{+T} \bar{P} \xi_s^+ - \xi_s^T \bar{P} \xi_s = \Delta V_s^1 + \Delta V_s^2 , \tag{A.17}$$

with $\Delta V_s^1 = x^{T+} P x^+ - x^T P x + \breve{x}^{T+} P \breve{x}^+ - \breve{x}^T P \breve{x}$ and $\Delta V_s^2 = tr\left\{\tilde{W}_u^{+T} \tilde{W}_u^+\right\} - tr\left\{\tilde{W}_u^T \tilde{W}_u\right\}$.

Evaluating along the reset dynamics (34) the first difference of the first part becomes

$$\Delta V_s^1 = x^T P x - \breve{x}^T P \breve{x}, \text{ for } t = \tau_k . \tag{A.18}$$

Now, using Frobenius norm we can express ΔV_s^1 as

$$\Delta V_s^1 \le -\lambda_{\min}\left(P\right)\left\|\breve{x}\right\|^2 + B_s^{jp}, \quad t = \tau_k . \tag{A.19}$$

where $\lambda_{\min}(P)$ is the minimum eigenvalue of the symmetric positive definite matrix P and $B_s^{jp} = \|P\|(B_1^x)^2$ where B_1^x is the bound for system state x during flow.

The first difference of the second Lyapunov function candidate ΔV_s^2 , it can be written from (A.8) in Lemma 2.

Finally, combining (A.19) and (A.8)

$$\Delta V_s\left(\xi_s\right) \le -\lambda_{\min}\left(P\right)\left\|\breve{x}\right\|^2 - \frac{a_1}{2}\left\|\tilde{W}_u\right\|^2 + B_s^{jp} + B_1^{\tilde{W}_u},\ t = \tau_k . \tag{A.20}$$

From (A.20) Lyapunov, the first difference $\Delta V_s\left(\xi_s\right) < 0$ as long as $\left\|\breve{x}\right\| > \sqrt{(B_s^{jp} + B_1^{\tilde{W}_u})/\lambda_{\min}\left(P\right)} = B_{s,2}^{\breve{x}}$ or $\left\|\tilde{W}_u\right\| > \sqrt{2(B_s^{jp} + B_1^{\tilde{W}_u})/a_1} \equiv B_{s,2}^{\tilde{W}_u}$. Since, $B_1^x \to B_{s,\max}^x$ and $B_1^{\tilde{W}_u} \to B_{\max}^{\tilde{W}_u}$ for all $\tau_k > \bar{T}$ from Case I and Lemma 2, respectively, the bounds $B_{s,2}^{\breve{x}} \to B_{s,\max}^{\breve{x}}$ and $B_{s,2}^{\tilde{W}_u} \to B_{s,\max}^{\tilde{W}_u}$ for all $\tau_k > \bar{T}$ where $B_{s,\max}^{\breve{x}} = \sqrt{(B_{s,\max}^{jp} + B_{\max}^{\tilde{W}_u})/\lambda_{\min}\left(P\right)}$ and $B_{s,\max}^{\tilde{W}_u} = \sqrt{(B_{s,\max}^{jp} + B_{\max}^{\tilde{W}_u})/a_1}$ with $B_{s,\max}^{jp} = \|P\| B_{s,\max}^{x^2}$. This implies, all the system variables x , $\breve{x}$, $\tilde{W}_u$ are ultimately bounded during the jump for all $\tau_k > \bar{T}$ or alternatively for all $t > T$.

Remark A.2: From both parts of the proof, the stability conditions (4) and (5) in Lemma 1 holds with the bound $\mu_s = \max\left\{B_{s,\max}^x, B_{s,\max}^{\breve{x}}, B_{s,\max}^{\tilde{W}}\right\}$. Further, since $\Delta V_s\left(\xi_s\right) < 0$, for $\xi_s > \mu_s$, $\xi_s^+ = \xi_s + F_d^s\left(\xi_s\right)$ is also bounded. Therefore, $\theta = \sup_{\xi_s \in \bar{\mathcal{B}}_{\mu_s} \cap \mathcal{Z}_s} V_s\left(\xi_s + F_d^s\left(\xi_s\right)\right)$ exists. Hence, we conclude that the augmented system state ξ_s of the impulsive dynamical

system (33) and (34) is locally ultimately bounded for all time $t > T$. To compute the ultimate bound, consider the Lyapunov equation (A.9). It is clear that

$$\lambda_{\min}\left(\bar{P}\right)\left\|\xi_s\right\|^2 \le \xi_s^T \bar{P} \xi_s \le \lambda_{\max}\left(\bar{P}\right)\left\|\xi_s\right\|^2 . \tag{A.21}$$

Therefore, the ultimate bound $\Xi = \sqrt{\eta / \lambda_{\min}\left(\bar{P}\right)}$ where $\eta = \max\left\{\lambda_{\max}\left(\bar{P}\right)\mu_s^2, \theta\right\}$. ■

Proof of Theorem 3: From the closed-loop dynamics (26) and the NN weight update law (19) the following inequality

$$\left\|\dot{x}\right\| \le \left\|A\right\|\left\|x\right\| + \nu_{1,k}, \quad \text{for} \quad \tau_k < t \le \tau_{k+1} \tag{A.22}$$

holds where $\nu_{1,k} = g_{\max}\left(\left\|\tilde{W}_{u,k}\right\|\varphi_{u,\max} + \varepsilon_{u,\max}\right) + 2g_{\max}L_{\varphi_u}\varphi_{u,\max}\left\|\hat{W}_{u,k}\right\|$ is a piece wise constant function since the NN weight estimation error $\tilde{W}_{u,k}$ and weight estimate $\hat{W}_{u,k}$ are constant for each k^{th} flow interval due to no weight update.

Consider the event-trigger error e_s. The derivative of e_s can be expressed as

$$de_s/dt \le \left\|\dot{e}_s\right\| = \left\|\dot{x} - \dot{\hat{x}}\right\| = \left\|\dot{x}\right\| \le \left\|A\right\|\left\|x\right\| + \nu_{1,k}, \tag{A.23}$$

for $\tau_k < t \le \tau_{k+1}$. By comparison lemma [23], the solution of the differential inequality (A.23) with initial condition $e_s^+ = 0$ for $t = \tau_k$ is upper bounded by

$$\left\|e_s\right\| \le \int_{\tau_k^+}^{t} \exp\left(\left\|A\right\|\left(t-s\right)\right)\nu_{1,k}\,ds$$

$$= \left(\nu_{1,k}/\left\|A\right\|\right)\left[\exp\left(\left\|A\right\|\left(t-\tau_k\right)\right) - 1\right], \quad \tau_k < t \le \tau_{k+1}. \tag{A.24}$$

The lower bound on k^{th} inter-event time, i.e., $\delta\tau_k = \tau_{k+1} - \tau_k$, is the time it takes $\left\|e_s\right\|$ to grow from 0 to the minimum value of the threshold $\sigma_{s,\min} = \min_k\left(\sigma_{s,k}\left\|x\right\|\right)$ over all flow interval. Note that the threshold coefficient σ_s in (36) is a piece wise constant

function since the NN weight, $\hat{W}_u$, is not updated during the flow period. Hence, the minimum value of the threshold, $\sigma_s \|x\|$, from (35) over flow interval $(\tau_k, \tau_{k+1}]$ for all $k = 1, 2, \cdots$ becomes

$$\sigma_{s,\min} B_{s,\max}^x = \left(\Gamma_s q_{\min} / 4 g_{\max} L_{\varphi_u} \hat{W}_{u,\max} \|P\| \right) B_{s,\max}^x, \tag{A.25}$$

where $\|x\| \geq B_{s,\max}^x$ for an event to trigger as per Remark 3 and $\hat{W}_{u,\max} = \max_k(\hat{W}_{u,k})$ is the maximum value of the NN weight estimates for all $k = 1, 2, \cdots$. The maximum value $\hat{W}_{u,\max}$ exists since the weight estimates are bounded for all $t > T$ and proven in Theorem 2.

Further, at the next event, i.e., τ_{k+1}, it holds that $\left\| e_s(\tau_{k+1}) \right\| = \sigma_{s,\min} B_{s,\max}^x$. Using this relation and comparing with (A.24), we arrive at

$$\sigma_{s,\min} B_{s,\max}^x \leq \left(\nu_{1,k} / \|A\| \right) \left(\exp\left(\|A\| (\tau_{k+1} - \tau_k) \right) - 1 \right). \tag{A.26}$$

Solving the inequality (A.26), the lower bound on inter-event times

$$\delta\tau_k \geq \left(1/\|A\| \right) \ln\left(1 + \left(\|A\| / \nu_{1,k} \right) \sigma_{s,\min} B_{s,\max}^x \right) \tag{A.27}$$

From (A.27) the lower bound on the inter-event times $\delta\tau_k > 0$ for all $k = 1, 2, \cdots$ since $\left(\|A\| / \nu_{1,k} \right) \sigma_{s,\min} B_{s,\max}^x > 0$. ■

Proof of Proposition 1: Recalling the definition of $\nu_{1,k}$ form (A.22) and using the definition $\left\| \hat{W}_u \right\| = \left\| W^* - \tilde{W}_u \right\| \leq \|W\| + \left\| \tilde{W}_u \right\|$ yields

$$\begin{aligned} \nu_{1,k} &\leq g_{\max} \left(\varphi_{u,\max} \left\| \tilde{W}_{u,k} \right\| + \varepsilon_{u,\max} \right) + 2 g_{\max} L_{\varphi_u} \varphi_{u,\max} \left(\|W_u\| + \left\| \tilde{W}_{u,k} \right\| \right) \\ &= g_{\max} \left(\varphi_{u,\max} \left(1 + 2 L_{\varphi_u} \right) \left\| \tilde{W}_{u,k} \right\| + \varepsilon_{u,\max} \right) + 2 g_{\max} L_{\varphi_u} \varphi_{u,\max} W_{u,\max} \equiv \nu_{M,k} \end{aligned} \tag{A.28}$$

Substituting (A.28) in (A.27), it holds that the lower bound on the inter-event times satisfies the relation

$$\delta\tau_k \geq (1/\|A\|)\ln(1+(\|A\|/\nu_{M,k})\sigma_{s,\min}B^x_{s,\max}).$$ ■

V. APPROXIMATE OPTIMAL CONTROL OF AFFINE NONLINEAR CONTINUOUS TIME SYSTEMS USING EVENT SAMPPLED NEURO DYNAMIC PROGRAMMING

Avimanyu Sahoo, Hao Xu and S. Jagannathan

Abstract *— This paper presents an approximate optimal control of nonlinear continuous-time systems in affine form by using the adaptive dynamic programming (ADP) with event sampled state and input vectors. The knowledge of the system dynamics are relaxed by using a neural network (NN) identifier with event sampled inputs. An approximate solution to the Hamilton-Jacobi-Bellman (HJB) equation, by using event sampled NN approximator, is introduced. Subsequently, the NN identifier and approximated value function are utilized to generate the optimal control policy. Both the identifier and value function weights are tuned only at the event sampled instants leading to an aperiodic update scheme. A novel adaptive event sampling condition is designed to determine the sampling instants such that the approximation accuracy and stability are maintained. A positive lower bound on the minimum inter-sample time is guaranteed to avoid accumulation point and the dependence of inter-sample time upon the NN weight estimate is analyzed in detail. The extension of Lyapunov theory is utilized to guarantee the local ultimate boundedness of the resulting nonlinear impulsive dynamical closed-loop system. Finally, a numerical example is utilized to evaluate the performance of the near optimal design through simulation studies. The net result is the design of event sampled ADP-based controller for nonlinear continuous-time systems.*

Index Terms **-** Adaptive dynamic programming, Hamilton-Jacobi-Bellman equation, event sampled control, neural networks, optimal control.

1. INTRODUCTION

Optimal control [1]-[2] of dynamic systems in continuous-time is a challenging problem due to the difficulty involved in obtaining a closed-form solution to the Hamilton-Jacobi-Bellman (HJB) [2] equation. Adaptive dynamic programming (ADP) [1]-[13] techniques, on the other hand, are used to solve the optimal control of uncertain systems online by finding an approximate value function which becomes a solution to HJB equation. Among the earlier works on ADP-based optimal control [1]-[13], the reinforcement learning technique is combined with dynamic programming, using an actor-critic neural network (NN) based framework [4], to generate an online yet approximate solution to the optimal control without needing the knowledge of system dynamics. Later, online policy iteration schemes [4] are introduced to obtain the solution of HJB equation and attain optimality. In addition, an alternate single NN-based ADP approach is presented in [10] for an affine nonlinear continuous-time system without using an iterative technique. The NN weights are tuned online and periodically to achieve near optimality.

For controller implementation, the traditional sampled data approach with a fixed sampling interval is found to be computationally intensive. Event-based sampling [14]-[15] and control, on the other hand, is increasingly gaining prominence among control researchers because of its computational and/or communication resource saving capability. In an event sampled framework, the state vector is sampled based on certain state dependent criteria referred to as event-triggering condition. The controller is executed at these aperiodic sampling instants. The event-triggering condition is designed by taking into account the stability and closed-loop performance, and, hence, proven to

be advantageous [14] over its periodic counterpart. However, the majority of the event-triggered techniques [14]-[15], are designed for stabilization without any performance criterion and under the assumption that the system dynamics are known.

In practice, the system dynamics may not be known accurately for the traditional backward-in-time optimal techniques to work. Therefore, development of an ADP scheme in an event sampled context is necessary and this is a challenging and an open problem. Motivated by the above limitations, in this paper, we propose a novel ADP method to solve the approximate optimal control of nonlinear continuous-time systems in an event sampled paradigm. After revisiting the NN approximation under this paradigm, NNs are subsequently used to identify the unknown system dynamics, and the value function which becomes the solution of the HJB equation. The optimal control policy is derived by using both the approximated system dynamics and the value function.

Although this work is motivated by [10] where a continuous-time near optimal controller is designed using a single NN as a value function approximator, this effort develops an event sampled NN approximation scheme to achieve near optimality. Another major difference with [10] is the aperiodic tuning of the NN weights and execution of the control at the event sampled instants. Above all the hybrid/impulse system [19]-[21] framework is used to analyze the stability due to aperiodic availability of state and input vector.

An adaptive sampling condition using actual NN weight estimates is analytically derived via Lyapunov techniques. Since, the actual NN weight estimates are tuned at the event sampled instants, the computation is reduced when compared to traditional NN based schemes [10], [16]. Next, the closed-loop system is formulated as a nonlinear

impulsive dynamical system [19]-[21] and the extension of Lyapunov direct method [19] is used to prove the local ultimate boundedness (UB) of all signals. A minimum inter-sample time is guaranteed to avoid accumulation point. It is demonstrated that the events will occur more frequently during the initial learning phase to attain the approximation accuracy. Nevertheless, the overall computational load is reduced over its traditional periodic sampled counterpart. A preliminary version of the paper is published as [22].

The remaining of the paper is organized as follows. Section 2 presents a brief background on the traditional ADP schemes and formulates the problem. Section 3 details the design procedure of the proposed event sampled ADP. The stability of the closed-loop system is analyzed in Section 4 followed by the simulation results in Section 5. Section 6 presents the conclusions. The detailed proof for the theorems and the lemmas are provided in the Appendix.

2. BACKGROUND AND PROBLEM STATEMENT

In this section, after introducing notations, a brief background of traditional ADP is presented. Then, the problem of event-sampled ADP is formulated.

2.1 NOTATIONS

Let $\Re^n$ is the n dimensional Euclidean space. For a vector $x(t) \in \Re^n$, we denote $\|x\|$ its vector 2-norm. For a matrix $A \in \Re^{n\times m}$, $\|A\|$ is its Frobenius norm, $A^T \in \Re^{m\times n}$ is the transpose of A and $vec(A)$ is the vector operator to stack the columns in a vector. For a square matrix $A \in \Re^{n\times n}$, $\lambda_{\max}(A)$ and $\lambda_{\min}(A)$ are the maximum and minimum eigenvalues of the matrix, and $tr\{A\}$ is the trace of A.

Consider the impulsive dynamical system [19] given by

$$\dot{\zeta} = F_c(\zeta),\ \zeta \in \mathcal{C},\ \zeta \notin \mathcal{Z},\ \zeta(0) = \zeta_0, \tag{1}$$

$$\Delta\zeta = F_d(\zeta);\ \zeta \in \Omega_\zeta,\ \zeta \in \mathcal{Z}, \tag{2}$$

where $\zeta \in \Omega_\zeta \subset \Re^{n_\zeta}$ represents the state vector, $F_c : \Omega_\zeta \to \Re^{n_\zeta}$ and $F_d : \Omega_\zeta \to \Re^{n_\zeta}$ denotes the nonlinear continuous and reset/jump dynamics, respectively. The set $\mathcal{C} \subset \Omega_\zeta$ and $\mathcal{Z} \subset \Omega_\zeta$ are respectively the flow and the jump sets and Ω_ζ is an open set with $0 \in \Omega_\zeta$. The difference is defined as $\Delta\zeta(t) = \zeta(t^+) - \zeta(t)$ where $\zeta(t^+) = \lim_{\varepsilon\to 0}\zeta(t+\varepsilon)$. For brevity, we write $\zeta\left(t^+\right) = \zeta^+$. The time variable t is dropped from the states and other functions where there is no ambiguity.

2.2 BACKGROUND

Consider the controllable nonlinear continuous-time system in affine form represented as

$$\dot{x}(t) = f(x(t)) + g(x(t))u(t), \quad x(0) = x_0 \quad , \tag{3}$$

where $x \in \Re^n$ and $u \in \Re^m$ are the state and the control input vectors, respectively. The unknown functions $f(x) \in \Re^n$ and $g(x) \in \Re^{n \times m}$ represent the nonlinear system dynamics satisfying $f(0) = 0$ with the following assumption.

Assumption 1[10]: The nonlinear system is controllable and observable. The nonlinear matrix function $g(x)$ for all $x \in \Omega_x$ satisfies $g_m \le \|g(x)\| \le g_M$, with g_M and g_m are known positive constants, and Ω_x is a compact set.

Consider the value function

$$V(t) = \int_t^\infty r(x(\tau), u(\tau))d\tau \, , \tag{4}$$

to be optimized where $r(x,u) = Q(x) + u^T R u$ is the cost-to-go function. The function $Q(x) \in \Re$ and the matrix $R \in \Re^{m \times m}$ are positive definite quadratic function and matrix, respectively, to penalize the system state vector and the control input. The initial control input u_0 must be admissible to keep the infinite horizon value function (4) finite.

The Hamiltonian for the cost function (4) can be given by

$$H(x,u) = Q(x) + u^T R u + \nabla_x V^T [f(x) + g(x)u],$$

where $\nabla_x V = \partial V / \partial x$ is the gradient with respect to x.

The optimal control policy $u^*(x)$ which minimizes the value function (4) can be computed using stationary condition as

$$u^*(x) = -(1/2) R^{-1} g^T(x) \left(\partial V^* / \partial x \right), \tag{5}$$

where $V^* \in \Re$ is the optimal value function. Then, substituting the optimal control input into the Hamiltonian, the HJB equation becomes

$$H^*(x, u^*) = Q(x) + \nabla_x V^*(x)\, f(x) - (1/4) \nabla_x V^{*T}(x)\, g(x) R^{-1} g^T(x) \nabla_x V^*(x) = 0, \tag{6}$$

where $\nabla_x V^*(x) = \partial V^* / \partial x$. It is extremely difficult to obtain an analytical solution to the HJB equation (6). Therefore, ADP based techniques [10] are utilized to generate an approximate solution in a forward-in-time manner by using periodically sampled state vector. Next, the problem for the event sampled ADP is formulated.

2.3 PROBLEM STATEMENT

In this section, we will formulate the event sampled optimal control problem by highlighting the challenges involved in the design with respect to approximation and stability. In an event sampled framework [14]-[15], the system state vector is sensed continuously and released to the controller only at the event-sampled instants.

To denote the sampled instants we define an increasing sequence of time instants $\{t_k\}_{k=1}^{\infty}$, referred to as event-sampled instants satisfying $t_{k+1} > t_k$, $\forall k = 1, 2, \cdots$ and $t_0 = 0$ is the initial sampling instant. The sampled state, $x(t_k)$, is released to the controller and the last sampled state at the controller denoted by $\breve{x}(t)$ is updated. It can be represented as a jump in the state $\breve{x}(t)$ at the event sampled instants and defined as

$$\breve{x}(t^+) = x\left(t_k\right),\ t = t_k,\ \forall k = 1, 2, \cdots. \tag{7}$$

Then it is held at the controller until the next update and is given by

$$\breve{x}(t) = x\left(t_k\right),\ t_k < t \le t_{k+1},\ \forall k = 1, 2, \cdots. \tag{8}$$

The error introduced due to the event-sampled state can be written as

$$e_{ET}(t) = x(t) - \breve{x}(t), t_k < t \le t_{k+1}, \forall k = 1, 2, \cdots, \tag{9}$$

where $e_{ET}(t)$ is referred to as event sampling error. Thus, the event sampling error is reset to zero with the update in the state i.e., $e_{ET}^{+} = 0$, $t = t_k$, $\forall k = 1, 2, \cdots$.

For optimal policy generation using ADP in an event sampled framework, the value function and the system dynamics need to be approximated with intermittently available system state vector. Therefore, to ensure desired accuracy of approximation, the universal approximation property of the NNs is revisited in the next few paragraphs.

By the universal approximation property of the NN a continuous function $h(x) \in \Re^n$ for all $x \in \Omega_x \subset \Re^n$ can be approximated as

$$h(x) = W_h^T \varphi_h(x) + \varepsilon_h(x) \ . \tag{10}$$

where $W_h \in \Re^{l \times n}$ is the unknown constant target NN weight matrix. The function $\varphi_h(x)$ is a bounded activation function, and $\varepsilon_h(x)$ is the traditional NN reconstruction error with l being the number of hidden-layer neurons. The implicit assumption here is that the state vector is available continuously for approximation.

For approximating the function $h(x)$ with an event sampled state vector $\breve{x}(t)$ defined in (7) and (8), the equation (A.1) can be rewritten as

$$\begin{aligned} h(x) &= W_h^T \varphi_h(x) - W_h^T \varphi_h(\breve{x}) + W_h^T \varphi_h(\breve{x}) + \varepsilon_h(x) \\ &= W_h^T \varphi_h(\breve{x}) + \varepsilon_{e,h}(\breve{x}, e_{ET}), \ t_k < t \le t_{k+1}, \forall k = 1, 2, \cdots, \end{aligned} \tag{11}$$

where $\varphi(\breve{x})$ is the event sampled activation function and $\varepsilon_{e,h}(\breve{x}, e_{ET}) = W_h^T [\varphi_h(\breve{x} + e_{ET}) - \varphi_h(\breve{x})] + \varepsilon_h(\breve{x} + e_{ET})$ is the event sampled reconstruction error. The reconstruction error

$\varepsilon_{e,h}(\breve{x}, e_{ET})$ consists of a second term $W_h^T[\varphi_h(\breve{x}+e_{ET})-\varphi_h(\breve{x})]$ in addition to the traditional reconstruction error, $\varepsilon_h(x)=\varepsilon_h(\breve{x}+e_{ET})$ which appears to be a function of the event sampling error, e_{ET}. It is clear that the event sampled approximation is equal to the traditional universal approximation if the event sampling error e_{ET} is zero. Since, our objective is to reduce computation by allowing this error to increase without affecting the stability, a tradeoff exists between the approximation accuracy and reduction in computation which is decided by the sampling frequency.

The optimal value function with event sampled state vector (11) can be written as

$$V^*(x)=W_V^T\phi(\breve{x})+\varepsilon_{e,V}(\breve{x}, e_{ET}) \ , \ t_k < t \le t_{k+1}, \tag{12}$$

where $W_V \in \Re^{l_V}$ is the unknown constant target NN weights, $\phi(\breve{x}) \in \Re^{l_V}$ is the event sampled activation function, $\varepsilon_{e,V}(\breve{x}, e_{ET})=W_V^T(\phi(x)-\phi(\breve{x}))+\varepsilon_V(\breve{x}+e_{ET})$ is the event-based reconstruction error with $\varepsilon_V\left(\breve{x}+e_{ET}\right)=\varepsilon_V\left(x\right)\in\Re$ is the traditional reconstruction error.

The HJB equation (6) with event sampled approximation of the value function (12) can be expressed as

$$\begin{aligned} &H^*(x,u^*)=Q(x)+\left(W_V^T\nabla_x\phi(\breve{x})+\nabla_x\varepsilon_{e,V}(\breve{x}, e_{ET})\right)f(x) \\ &-(1/4)\left(W_V^T\nabla_x\phi(\breve{x})+\nabla_x\varepsilon_{e,V}(\breve{x}, e_{ET})\right)(x)\,D(x)\left(\nabla_x\phi^T(\breve{x})W_V+\nabla_x\varepsilon_{e,V}(\breve{x}, e_{ET})\right), \end{aligned} \tag{13}$$

where $D(x)=g(x)R^{-1}g^T(x)$, $\nabla_x\phi(\breve{x})=\partial\phi(\breve{x})/\partial x$, and $\nabla_x\varepsilon_{e,V}(\breve{x}, e_{ET})=\partial\nabla_x\varepsilon_{e,V}(\breve{x}, e_{ET})/\partial x$. It is clear from (13), the HJB equation is also a function of the event sampling error e_{ET}. In other words, the performance is governed by the event sampling condition design.

Thus, the event sampled optimal control problem can be defined more precisely as follows: 1) Approximate the unknown system dynamics $f(x)$ and $g(x)$, and the value function V in an event sampled context with a desired level of accuracy; 2) Design the event sampling condition not only to reduce computation but also to minimize approximation error. Finally, 3) guarantee a positive lower bound on the inter-sample time. A solution along with the detailed design procedure of the event sampled near optimal control design is presented in the next section.

3. EVENT SAMPLED NEAR OPTIMAL CONTROL DESIGN

The proposed structure of the event sampled near optimal design is illustrated in Figure 1 and the design will be carried out using two NNs with event sampled state vector. One NN is used as an identifier to approximate the unknown system dynamics and the second one is used to approximate the solution of the HJB equation which is the value function. Now to reduce the computation and ensure accuracy of the approximation, we propose an adaptive event sampling condition as the function of event sampling error, the estimated NN weights and the system state vector. The system state is sent to the controller at the event sampled instants and used to tune the NN weights. The weights are held during the inter-sample times and, hence, tuning scheme becomes aperiodic.

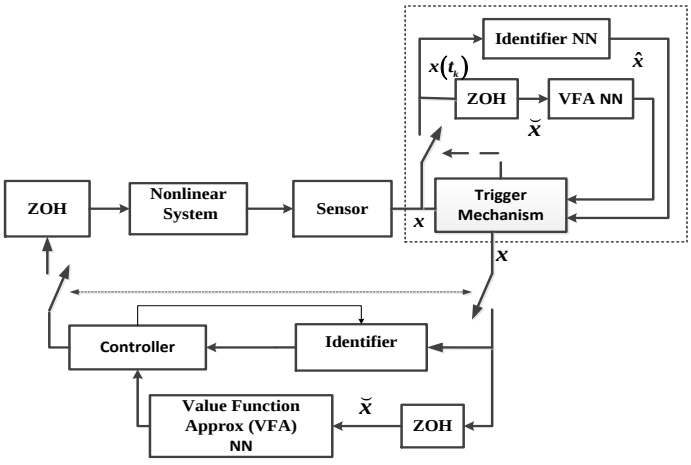

Figure 1 . Near optimal event-sampled control system

Remark 1: To evaluate the proposed adaptive event sampling condition at the trigger mechanism, in case of an NCS [54], will require transmission of the NN weight estimates from the controller. To mitigate this additional transmission cost, mirror identifier and

value function approximator NNs are used at the trigger mechanism to estimate the NN weights locally. Both the actual and mirror NNs operate in synchronism at the event sampled instants. Thus, this design can be considered as an event sampled NCS with negligible delays and packet losses.

The detailed design procedure of the NN identifier and the controller are presented along with their weight tuning laws in the next subsections.

3.1 IDENTIFIER DESIGN

The knowledge of the system dynamics, $f(x)$ and $g(x)$ are needed for the computation of the optimal control policy (5). To relax this, an event sampled NN based identifier design is presented in this subsection. By using the event-based approximation (11), the nonlinear continuous-time system in (1) can be represented as

$$\dot{x} = f(x) + g(x)u = W_I^T \sigma_I(\breve{x})\bar{u} + \varepsilon_{e,I}, \tag{14}$$

where $W_I = [W_f^T \;\; W_g^T]^T \in \Re^{(l_f + ml_g)\times n}$ with $W_f \in \Re^{l_f \times n}$ and $W_g \in \Re^{ml_g \times n}$ are the unknown target NN weight matrices, $\sigma_I(\breve{x}) = \begin{bmatrix} \sigma_f(\breve{x}) & 0 \\ 0 & \sigma_g(\breve{x}) \end{bmatrix}$ with $\sigma_f(\breve{x}) \in \Re^{l_f}$ and $\sigma_g(\breve{x}) \in \Re^{ml_g \times m}$ are the event sampled activation functions. The error $\varepsilon_{e,I} = \begin{bmatrix} \varepsilon_{e,f}(\breve{x}, e_{ET}) & \varepsilon_{e,g}(\breve{x}, e_{ET}) \end{bmatrix}\bar{u}$ is the event-based reconstruction error with $\varepsilon_{e,f}(\breve{x}, e_{ET}) = W_f^T(\sigma_f(x) - \sigma_f(\breve{x})) + \varepsilon_f(\breve{x} + e_{ET})$, $\varepsilon_{e,g}(\breve{x}, e_{ET}) = W_g^T\left(\sigma_g(x) - \sigma_g(\breve{x})\right) + \varepsilon_g(\breve{x} + e_{ET})$ and $\bar{u} = [1 \; u^T]^T$. The subscript f and g denotes parameters corresponding to the functions $f(\bullet)$ and $g(\bullet)$, respectively. The event sampled reconstruction error can also be written as $\varepsilon_{e,I} = W_I^T \tilde{\sigma}_I(x, \breve{x})\bar{u} + \varepsilon_I$ where $\tilde{\sigma}_I(x, \breve{x}) = \sigma_I(x) - \sigma_I(\breve{x})$ is the activation

function error and $\varepsilon_I = [\varepsilon_f \;\; \varepsilon_g u]$ the augmented traditional reconstruction error. The constants l_f and l_g denote the number of neurons of the NNs. The following assumption holds for the NN [16].

Assumption 2: The target weight vector, W_I, the activation function, $\sigma_I(\bullet)$, and the reconstruction error $\varepsilon_I(\bullet)$ are bounded above in a compact set such that $\|W_I\| \le W_{\mathrm{I},M}$, $\|\sigma_I(\bullet)\| \le \sigma_{I,M}$, $\|\varepsilon_I(\bullet)\| \le \varepsilon_{I,M}$, satisfied where $W_{I,M}$, $\sigma_{I,M}$, $\varepsilon_{I,M}$ are positive constants. Further, it is assumed that the activation function $\sigma_I(x)$ is Lipschitz continuous in a compact set and satisfies $\|\sigma_I(x) - \sigma_I(\breve{x})\| \le C_{\sigma_I} \|x - \breve{x}\|$ where $C_{\sigma_I} > 0$ is a constant.

Since the system state vector is only available at event-sampled instants, the event-based identifier dynamics is defined as

$$\dot{\hat{x}} = A(\hat{x} - \breve{x}) + \hat{f}(\breve{x}) + \hat{g}(\breve{x})u, \; t_k < t \le t_{k+1}, \tag{15}$$

where $\hat{x} \in \Re^n$ is the identifier estimated state vector, A is a user defined Hurwitz matrix and satisfies the Lyapunov equation $A^T P + PA = -\Pi$ where P and Π are positive definite matrices. The matrix A ensures the stability of the identifier. The functions $\hat{f}(\breve{x}) \in \Re^n$ and $\hat{g}(\breve{x}) \in \Re^{n \times m}$ are the estimated system dynamics. By using the event sampled NN approximation for the system as in (14), the estimated value of identifier dynamics are represented as

$$\dot{\hat{x}} = A(\hat{x} - \breve{x}) + \hat{W}_I^T(t)\sigma_I(\breve{x})\bar{u}, \quad t_k < t \le t_{k+1}, \tag{16}$$

where $\hat{W}_I(t)$ is the estimated NN weight matrices and $\sigma_I(\breve{x})$ being event-sampled the activation functions.

Defining $e_I = x - \hat{x}$ as the identification error, the identification error dynamics, from (14) and (16), can be expressed as

$$\dot{e}_I = Ae_I - Ae_{ET} + \tilde{W}_I^T \sigma_I(\breve{x})\bar{u} + \hat{W}_I^T \tilde{\sigma}_I(x, \breve{x})\bar{u} + \varepsilon_I . \tag{17}$$

Since, e_I can only be computed at the event-sampled instants with available current sampled state at the identifier, the NN identifier weight matrices are tuned at the event-sampled instants only. This can be considered as a jump in the identifier NN weights which is given by

$$\hat{W}_I^+ = \hat{W}_I + \left(\alpha_I \sigma_I(x)\bar{u}e_I^T \Big/ (c + \|\bar{u}\|^2 \|e_I\|^2)\right) - \alpha_I \hat{W}_I ,\ t = t_k , \tag{18}$$

where $\alpha_I > 0$ denotes the learning rate and $c > 0$ is a positive constant. During the inter-sample times referred as flow duration, $t_k < t \le t_{k+1}$, the weights are held at the previously tuned values. Therefore, the tuning law during the inter-sample times or flow duration, $t_k < t \le t_{k+1}$, is given by

$$\dot{\hat{W}}_I = 0 ,\ t_k < t \le t_{k+1} . \tag{19}$$

From (18) and (19), it is clear that the NN weights are tuned aperiodically and, hence, saves computation when compared to traditional NN [10], [16]. To ensure the ultimate boundedness of the closed-loop system parameters, the approximated control coefficient function is held once it becomes less than equal to the lower bound. It can be expressed as

$$\|\hat{g}(\bullet)\| = \begin{cases} g_{\min}, & \text{if } \|\hat{g}(x)\| \le g_{\min}, \\ \|\hat{g}(\bullet)\|, & \text{otherwise} . \end{cases} \tag{20}$$

Define the NN weight estimation error $\tilde{W}_I = W_I - \hat{W}_I$. The weight estimation error dynamics at both jump and flow duration can be expressed as

$$\tilde{W}_I^+ = \tilde{W}_I - \left(\alpha_I \sigma_I(x)\bar{u}e_I^T \Big/ (c + \|\bar{u}\|^2 \|e_I\|^2)\right) + \alpha_I \hat{W}_I,\ t = t_k, \tag{21}$$

$$\dot{\tilde{W}}_I = \dot{W}_I - \dot{\hat{W}}_I = 0,\ \ t_k < t < t_{k+1}. \tag{22}$$

We will use the identifier dynamics to design an event-sampled near optimal controller in the next subsection.

3.2 CONTROLLER DESIGN

In this section, the solution to the HJB equation, essentially the value function, is approximated by a second NN with event-sampled state vector. The approximated value function is utilized to obtain the near optimal control input. Consider the event sampled approximation of the optimal value function in (12). The following assumptions hold for the NN.

Assumption 3: The target NN weights, activation functions and the traditional reconstruction errors are bounded above satisfying $\|W_V\| \le W_{V,M}$, $\|\phi(\bullet)\| \le \phi_M$, and $\left|\varepsilon_V(\bullet)\right| \le \varepsilon_{V,M}$ with $W_{V,M}$, ϕ_M , and $\varepsilon_{V,M}$ are positive constants. It is further assumed that the gradient of the activation function and the reconstruction error are bounded by a positive constant, i.e., $\nabla_x \phi(\bullet) = \partial\phi(\bullet)/\partial x \le \phi_M^{'}$ and $\|\nabla_x \varepsilon_V(\bullet)\| = \partial\varepsilon_V(\bullet)/\partial x \le \varepsilon_{V,M}^{'}$.

Assumption 4: The activation function and its gradients are Lipschitz continuous in a compact set such that for $x \in \Omega_x \subset \Re^n$ and there exist positive constants $C_\phi > 0$ and $C_{\nabla\phi} > 0$ satisfying $\left\|\phi(x) - \phi(\breve{x})\right\| \le C_\phi \|x - \breve{x}\|$ and $\left\|\nabla_x \phi(x) - \nabla_x \phi(\breve{x})\right\| \le C_{\nabla\phi} \|x - \breve{x}\|$.

The optimal control policy (5) in terms of event-sampled NN approximation of the value function becomes

$$u^* = -(1/2)R^{-1}g^T(x)\left(\nabla_x \phi^T(\breve{x})W_V + \nabla_x \varepsilon_{e,V}(\breve{x}, e_{ET})\right). \quad (23)$$

The estimated value function in the context of event sampled state can be represented as

$$\hat{V}(\breve{x}) = \hat{W}_V^T \phi(\breve{x}),\ t_k < t \le t_{k+1}. \quad (24)$$

Therefore, the actual control policy by using the estimated value function (22) and the identifier dynamics is given by

$$u(x) = -(1/2)R^{-1}\hat{g}^T(\breve{x})\nabla_x \phi^T(\breve{x})\hat{W}_V = -(1/2)R^{-1}\left(\hat{W}_g^T \sigma_g(\breve{x})\right)^T \nabla_x \phi^T(\breve{x})\hat{W}_V,\ t_k < t \le t_{k+1}. \quad (25)$$

Now with the estimated value function (24) and approximated system dynamics (15), the error introduced in the HJB equation (13), referred to as temporal difference (TD) or HJB equation error, can be expressed as

$$\hat{H}(\breve{x}, u) = Q(\breve{x}) + u^T R u + \nabla_x \hat{V}^T(\breve{x})[\hat{f}(\breve{x}) + g(\breve{x})u], \quad (26)$$

for $t_k < t \le t_{k+1}$ where $\nabla_x \hat{V}(\breve{x}) = \partial \hat{V}(\breve{x}) / \partial x$. Substituting the actual control policy (25) in (26), the TD or HJB equation error can be expressed as

$$\hat{H}(\breve{x}, \hat{W}_V) = Q(\breve{x}) + \hat{W}_V^T \nabla_x \phi(\breve{x})\hat{f}(\breve{x}) - (1/4)\hat{W}_V^T \nabla_x \phi(\breve{x}) D(\breve{x}) \nabla_x^T \phi(\breve{x}) \hat{W}_V,\ t_k < t \le t_{k+1}, \quad (27)$$

where $\hat{D}(\breve{x}) = \hat{g}(\breve{x})R^{-1}\hat{g}^T(\breve{x})$.

Similar to the identifier, the value function is updated at the event-sampled instants with the updated HJB error. The HJB error (27) with event sampled state at the trigger instants, i.e., $\breve{x}^+ = x$, $t = t_k$ can be written as

$$\hat{H}^+(x, \hat{W}_V) = Q(x) + \hat{W}_V^T \nabla_x \phi(x)\hat{f}(x) - (1/4)\hat{W}_V^T \nabla_x \phi(x) D(x) \nabla_x^T \phi(x) \hat{W}_V,\ \ t = t_k. \quad (28)$$

The value function NN tuning law at the event sampled instants is selected as

$$\hat{W}_V^+ = \hat{W}_V - \left(\alpha_V \hat{\omega}\hat{H}^{+T}(x,\hat{W}_V)/(1+\hat{\omega}^T\hat{\omega})^2\right), t = t_k, \tag{29}$$

where $\alpha_V > 0$ is the NN learning gain parameter and

$$\hat{\omega} = \nabla_x\phi(x)\hat{f}(x) - (1/2)\nabla_x\phi(x)\hat{D}(x)\nabla_x^T\phi(x)\hat{W}_V. \tag{30}$$

During the inter-sample times or flow period, the tuning law for the value function NN is given as

$$\dot{\hat{W}}_V = 0, \;\; t_k < t \le t_{k+1}. \tag{31}$$

Define the value function NN weight estimation error as $\tilde{W}_V = W_V - \hat{W}_V$. The NN weight estimation error dynamics by using (29) and (31) can be expressed as

$$\tilde{W}_V^+ = \tilde{W}_V + \left(\alpha_V \hat{\omega}\hat{H}^{+T}(x,u)/(1+\hat{\omega}^T\hat{\omega})^2\right), \; t = t_k, \tag{32}$$

$$\dot{\tilde{W}}_V = 0, \;\; t_k < t \le t_{k+1}. \tag{33}$$

The HJB or TD error in terms of the value function NN weight estimation error $\tilde{W}_V$, using (13) and (27) can be expressed as

$$\begin{aligned}\hat{H}(\breve{x},\hat{W}_V) &= Q(\breve{x}) - Q(x) + \hat{W}_V^T\nabla_x\phi(\breve{x})\hat{f}(\breve{x}) - W_V^T\nabla_x\phi(x)f(x) \\ &-(1/4)\hat{W}_V^T\nabla_x\phi(\breve{x})D(\breve{x})\nabla_x^T\phi(\breve{x})\hat{W}_V - (1/4)W_V^T\nabla_x\phi(x)D(x)\nabla_x\phi^T(x)W_V - \varepsilon_H, \; t_k < t \le t_{k+1},\end{aligned} \tag{34}$$

where $\varepsilon_H = \nabla_x^T\varepsilon_V(x)\left(f(x) + g(x)u^*\right) - (1/4)\nabla_x^T\varepsilon_V(x)D(x)\nabla_x\varepsilon_V(x)$. It is routine to check that $\|\varepsilon_H\| \le \varepsilon_{H,M}$ where $\varepsilon_{H,M}$ is a positive constant.

Similarly, the HJB equation error at the event-sampled instants with $\breve{x}^+ = x$ can be computed from (34), and found to be

$$\begin{aligned}\hat{H}^+(x,\hat{W}_V) &= -\tilde{W}_V^T\nabla_x\phi(x)\hat{f}(x) - W_V^T\nabla_x\phi(x)\tilde{f}(x) + (1/2)\tilde{W}_V^T\nabla_x\phi(x)D(x)\nabla_x^T\phi(x)\hat{W}_V \\ &+(1/4)\tilde{W}_V^T\nabla_x\phi(x)D(x)\nabla_x^T\phi(x)\tilde{W}_V + (1/4)W_V^T\nabla_x\phi(x)D(x)\nabla_x\phi^T(x)W_V - \varepsilon_H, \;\; t = t_k.\end{aligned} \tag{35}$$

4. EVENT SAMPLING CONDITION AND MINIMUM INTER-SAMPLE TIME

In this section, we design an adaptive event sampling condition and present theoretical results. Before proceeding, the following stability notions are important.

Definition 1 [19]: The nonlinear state dependent impulsive dynamical system (1) and (2) is *ultimately bounded* with bound μ if there exists $\gamma > 0$ such that, for every $\delta \in (0, \gamma)$, there exists $T = T(\delta, \mu) > 0$ such that $\|\zeta(0)\| < \delta$ implies $\|\zeta(t)\| < \mu,\ t \geq T$.

4.1 IMPULSIVE DYNAMICAL SYSTEM

Consider the augmented state vector $\zeta = [x^T \ \ e_I^T \ \ \breve{x}^T \ \ vec(\tilde{W}_I)^T \ \ \tilde{W}_V^T]^T$. The dynamics during the flow $t_k < t \leq t_{k+1}$ and the jump instants, $t = t_k$, are computed as follows.

4.1.1 Flow Dynamics. The closed-loop system dynamics during the flow $t_k < t \leq t_{k+1}$, can be represented by using (1) and the control policy (25) as

$$\dot{x} = f(x) + g(x)\left(-(1/2)R^{-1}\hat{g}^T(\breve{x})\nabla_x\phi^T(\breve{x})\hat{W}_V\right). \tag{36}$$

Adding and subtracting $g(x)u^*$ in (36) and after some simple mathematical operations, the closed-loop dynamics during the flow can be written as

$$\begin{aligned}\dot{x} = {} & f(x) + g(x)u^* + (1/2)g(x)R^{-1}\tilde{g}^T(x)\nabla_x\phi^T(x)W_V + (1/2)g(x)R^{-1}\hat{g}^T(x)\nabla_x\phi^T(x)\tilde{W}_V \\ & + (1/2)g(x)R^{-1}\hat{g}^T(\breve{x})\left(\nabla_x\phi^T(x) - \nabla_x\phi^T(\breve{x})\right)\hat{W}_V + (1/2)g(x)R^{-1}g^T(x)\nabla_x\varepsilon_V(x),\end{aligned} \tag{37}$$

for $t_k < t \leq t_{k+1}$.

The flow dynamics for the identification error e_I is same as in (17). The last held state, $\breve{x}$, during the flow period remains constant. Thus the dynamics of the last held state, $\breve{x}$, is given by

$$\dot{\bar{x}} = 0, \quad t_k < t < t_{k+1}. \tag{38}$$

Finally, the dynamics of the NN weight estimation errors $vec(\tilde{W}_I)$ and $\tilde{W}_V$ during the flow are as in (22) and (33) represented in vector form.

Combining (17), (22), (33), (37), and (38), the flow dynamics for $t_k < t \le t_{k+1}$ of the impulsive dynamical system as

$$\dot{\zeta} = F_c(\zeta), \ \zeta \in \Omega_\zeta, \zeta \in \mathcal{C}, \zeta \notin \mathcal{Z}, \tag{39}$$

where $F_c = [F_c^{sT} \ \ F_c^{e_I T} 0^T \ \ 0^T \ \ 0^T]^T$ with $F_c^s = f(x) + g(x)u^* + (1/2)g(x)R^{-1}\tilde{g}^T(x)\nabla_x\phi^T(x)W_V$

$+(1/2)g(x)R^{-1}\hat{g}^T(x)\nabla_x\phi^T(x)\left(W_V - \tilde{W}_V\right) + (1/2)g(x)R^{-1}\hat{g}^T(\bar{x})\left(\nabla_x\phi^T(x) - \nabla_x\phi^T(\bar{x})\right)\left(W_V - \tilde{W}_V\right)$

$+(1/2)g(x)R^{-1}g^T(x)\nabla_x\varepsilon_V(x)$, $F_c^{e_I} = Ae_I - Ae_{ET} + \tilde{W}_I^T\sigma_I(\bar{x})\bar{u} + \hat{W}_I^T\tilde{\sigma}_I(x,\bar{x})\bar{u} + \varepsilon_I$ and 0^T s are null vectors of appropriate dimensions. The set $\mathcal{C} \subset \Omega_\zeta$ is the flow set and $\Omega_\zeta \subset \Re^{n_\zeta}$ a open set with $0 \in \Omega_\zeta$, $n_\zeta = n + n + l_I + l_V$ and $l_I = (l_f + ml_g)n$.

4.1.2 Jump Dynamics. The jump dynamics of the system state vector and the identification error are given by

$$x^+ = x, \quad t = t_k, \tag{40}$$

$$e_I^+ = x^+ - \hat{x}^+ = x - \hat{x} = e_I, \ \ t = t_k. \tag{41}$$

The jump dynamics of the last held state, $\bar{x}$, is given by

$$\bar{x}^+ = x, \quad t = t_k. \tag{42}$$

Further, the jumps in the NN weight estimation errors are given by (21) and (32).

Defining the first difference $\Delta\zeta = \zeta^+ - \zeta$ and using (21), (32), (40), (41), and (42), the difference equation for the reset/jump dynamics can be written as

$$\Delta\zeta = F_d(\zeta),\ \zeta \in \Omega_\zeta,\ \zeta \in \mathcal{Z}, \tag{43}$$

where $F_d(\zeta) = [0^T \ \ 0^T \ \ e_{ET}^T \ \ vec(\Delta\tilde{W}_I)^T \ \ (\Delta\tilde{W}_V)^T]^T$ with $\Delta\tilde{W}_I = -\frac{\alpha_I \sigma_I(x)\bar{u}e_I^T}{c + \|\bar{u}\|^2 \|e_I\|^2} + \alpha_I\left(W_I - \tilde{W}_I\right)$ and $\Delta\tilde{W}_V = \frac{\alpha_V \hat{\omega}\hat{H}^+(x,\hat{W}_V)}{(1+\hat{\omega}^T\hat{\omega})^2}$. The set $\mathcal{Z} \subset \Omega_\zeta$ the jump set. The flow and the jump set are decided by the event sampling condition introduced next.

Consider the event sampling error in (9). Define the event sampling condition given by

$$D(\|e_{ET}\|) \le \sigma_{ETC}\left(x, \|\hat{W}_V\|, \|\hat{W}_I\|\right), \tag{44}$$

where $\sigma_{ETC}\left(x, \|\hat{W}_V\|, \|\hat{W}_I\|\right) = \min\left\{\Gamma Q(x)/\eta_1, \sqrt{2\Gamma Q(x)/\eta_2}\right\}$ is the threshold with $\eta_1 = g_M \lambda_{\max}(R^{-1})\sigma_{I,M}\phi_M' \|\hat{W}_V\| \|\hat{W}_I\| C_{\nabla\phi}$, $\eta_2 = C_{\nabla\phi} g_M \lambda_{\max}(R^{-1})\sigma_{I,M}\phi_M' \|\hat{W}_I\| + (32/\Pi_{\min})\left(C_{\sigma_I}\right.$ $\left.\times \|\hat{W}_I\|^2 \|\bar{u}\|^2 \|P\|^2 + \|AP\|^2\right)$. The constants $0 < \Gamma < 1$, $C_{\nabla\phi}$ and C_{σ_I} are the Lipschitz constants and $\Pi_{\min} = \lambda_{\min}(\Pi)$. To ensure the estimated NN weights $\|\hat{W}_V\|$ and $\|\hat{W}_I\|$ in (44) are non-zero while executing the event sampling condition, the previous non zero update of the NN weight estimates are used to evaluate the event sampling condition when the updates become zero. The dead-zone operator $D(\bullet)$ is defined as

$$D(\bullet) = \begin{cases} \bullet, & \|x\| > B_{ub}^x, \\ 0, \end{cases} \tag{45}$$

where B_{ub}^x is the desired ultimate bound for the closed-loop system. The event sampled instants are decided at the violation of the condition (44).

Remark 2: The main advantage of this event sampling condition is that its threshold gets updated with NN weight estimates and system state. Therefore, states will be sampled based on the NN weight estimation errors and the system performance. This ensures the accuracy of approximation. Once the NN weights converged close to their target values and becomes constant, the threshold becomes similar to those used in traditional event sampling condition [14]-[15].

Remark 3: The dead zone operator $D(\bullet)$ prevents unnecessary triggering due to the NN reconstruction error once the system state is inside the ultimate bound.

To show the locally ultimately boundedness of the closed-loop event-sampled system we will use the following lemma as in [19] for the nonlinear impulsive dynamical systems. Before claiming the main results the following technical results are necessary.

Lemma 1: Consider the definition of $\hat{\omega}$ given in (30). For any positive number $N > 0$, the following relation holds

$$
\begin{aligned}
&-\frac{\tilde{W}_V^T \hat{\omega}\hat{\omega}^T \tilde{W}_V}{(1+\hat{\omega}^T\hat{\omega})^2} \le \frac{1}{N(1+\hat{\omega}^T\hat{\omega})^2}\tilde{W}_V^T\left((1/2)\nabla_x\phi(x)\hat{D}(x)\nabla_x^T\phi(x)W_V - \nabla_x\phi(x)\hat{f}(x)\right)\\
&\left((1/2)\nabla_x\phi(x)\hat{D}(x)\nabla_x^T\phi(x)W_V - \nabla_x\phi(x)\hat{f}(x)\right)^T\tilde{W}_V \qquad (46)\\
&-\frac{1}{4(N+1)(1+\hat{\omega}^T\hat{\omega})^2}(\tilde{W}_V^T\nabla_x\phi(x)\hat{D}(x)\nabla_x^T\phi(x)\tilde{W}_V)(\tilde{W}_V^T\nabla_x\phi(x)\hat{D}(x)\nabla_x^T\phi(x)\tilde{W}_V)^T.
\end{aligned}
$$

Proof: Refer to Appendix.

We will use the above results to show the locally ultimately boundedness of the NN weight estimation errors in the following lemma.

Lemma 2: Consider the nonlinear continuous-time system (3) along with the NN based identifier (16) and the value function approximator (24) with event sampled state vector. Let the Assumptions 1 through 4 hold and the initial identifier and value function NN

weights, $\hat{W}_I(0)$ and $\hat{W}_V(0)$, respectively, be initialized with nonzero finite values in the compact sets Ω_{W_I} and Ω_{W_V}. Suppose there exist a positive minimum inter-sample time between two consecutive event sampling instants, $\tau_{\min} = t_{k+1} - t_k > 0$, and the value function activation function $\phi(x)$ satisfies the persistency of excitation condition. Then, the weight estimation errors $\tilde{W}_I$ and $\tilde{W}_V$ are locally ultimately bounded (UB) for all event sampling instant $t_k > \bar{T}$ or $t > T$ for $T > \bar{T}$ provided the NN weights are tuned by using (18) and (19), and the learning gains satisfies $0 < \alpha_I < 1/3$, $0 < \alpha_V < \min\left\{\frac{1}{40}, \frac{2N(1-N)}{2(N+1)(16+N)}\right\}$ with $0 < N < 1$.

Proof: Refer to the Appendix.

Theorem 1: Consider the nonlinear continuous-time system (3), identifier (16), and the value function approximator (24) represented as an impulsive dynamical system (39) and (43). Let u_0 be an initial stabilizing control policy for (3). Let the Assumptions 1 through 3 hold, and assume there exist a minimum inter-sample time $\tau_{\min} > 0$. Suppose the value function activation function $\phi(x)$ satisfies the persistency of excitation condition and the system states are transmitted at the violation of the event-trigger condition (44). Let the initial identifier and value function NN initial weights, $\hat{W}_I(0)$ and $\hat{W}_V(0)$, respectively, are nonzero and bounded in compact sets Ω_{W_I} and Ω_{W_V}, and updated according to (29) and (31). Then, the closed-loop impulsive dynamical system is locally UB for all event sampling instant $t_k > \bar{T}$ or $t > T$ for $T > \bar{T}$. Further, the estimated value function

satisfies $\left\|V^* - \hat{V}\right\| \le B_V$ with B_V is small positive constants provided the design parameters are selected as in Lemma 2.

Proof: Refer to the Appendix.

The assumption on the inter-sample time in Theorem 1 is relaxed by guaranteeing the existence of a positive minimum inter-sample time the next subsection. The flow chart in Figure 2 illustrates the implementation of the event sampled ADP scheme designed in Section 3 and 4.

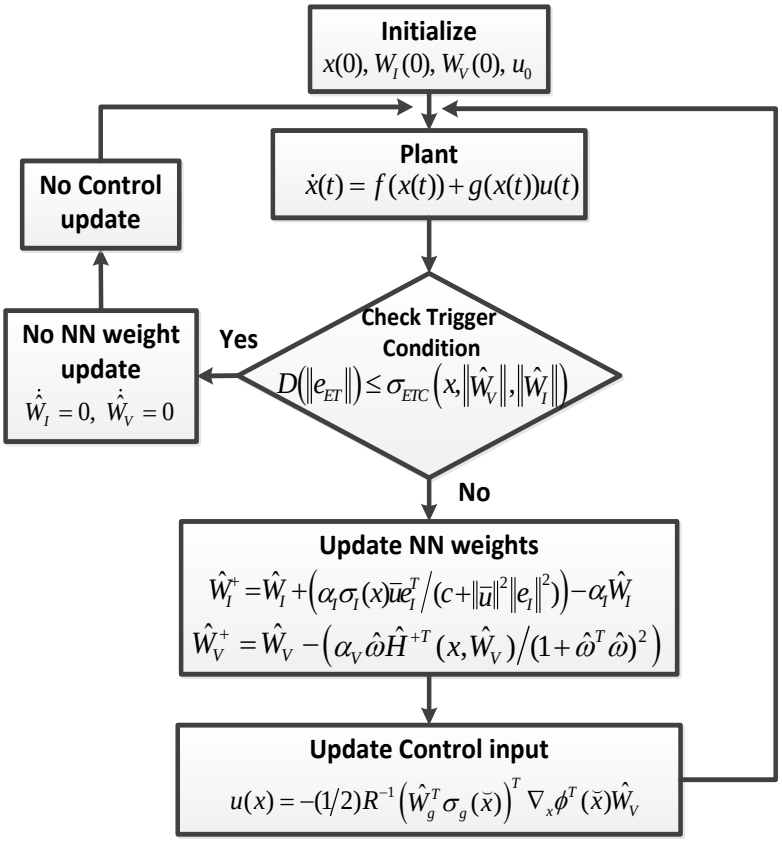

Figure 2. Flowchart of event sampled ADP scheme.

4.2 MINIMUM INTER-SAMPLE TIME

The following theorem guarantees the existence of the non-zero positive minimum inter-sample time $\tau_{\min} = \min_{k\in\mathbb{N}}\{t_{k+1} - t_k\}$.

Theorem 2: Consider the event sampled continuous time system (3) represented as impulsive dynamical systems (39) and (43) along with the event sampling condition (44). Then, the minimum inter-sample time $\tau_{\min}$, implicitly defined by (44), is lower bounded by a nonzero positive constant and it is given by

$$\tau_{\min} \geq \min_{k\in\mathbb{N}}\left\{(1/K)\ln\left(1+\left(K/M_k\right)\sigma_{ETC,\min}\right)\right\} > 0, \tag{47}$$

where $\sigma_{ETC,\min}$ is the minimum threshold coefficient value, $K > 0$ is a constant and

$M_k = g_M \lambda_{\max}(R^{-1})\sigma_{I,M}\phi_M^{'}\left\|\hat{W}_{I,k}\right\|\left\|\hat{W}_{V,k}\right\| + (1/2)g_M\lambda_{\max}(R^{-1})\phi_M^{'}W_{V,M}\left(\sigma_{I,M}\left\|\tilde{W}_{I,k}\right\| + \bar{\varepsilon}_I\right)$

$+(1/2)g_M\lambda_{\max}(R^{-1})\sigma_{I,M}\phi_M^{'}\left\|\hat{W}_{I,k}\right\|\left\|\tilde{W}_{V,k}\right\| + (1/2)g_M^2\lambda_{\max}(R^{-1})\varepsilon_{V,M}^{'}$ where the subscript k represents the k^{th} inter-sample time.

Proof. Refer to the Appendix.

Remark 4: The constant K satisfies the inequality $\left\|f(x) + g(x)u^*\right\| \leq K\|x\|$. This inequality holds [22] since the optimal control input is stabilizing.

Remark 5: It is clear from (47) that the lower bound on inter-sample times depends on M_k or alternatively, on the NN weight estimation errors $\tilde{W}_V$, and $\tilde{W}_I$ by the definitions $\hat{W}_I = W - \tilde{W}_I$ and $\hat{W}_V = W - \tilde{W}_V$. During the initial learning phase of the NN, the term M_k in (47) may become large for certain initial values $\hat{W}_V(0)$ and $\hat{W}_I(0)$, which may lead to smaller inter-sample times. Hence, a proper initialization of the NN weights, $\hat{W}_V(0)$ and

$\hat{W}_I(0)$ is necessary during learning phase. In addition, with update in NN weights, the convergence of the NN weight estimation errors will elongate the inter-sample times leading to fewer sampled events and less resource utilization.

5. SIMULATION RESULTS

In this section, the dynamics of a two-link robot manipulator is considered for simulation. The dynamics are given in by

$$\dot{x} = f(x) + g(x)u, \tag{48}$$

where $f(x) = \begin{bmatrix} x_3 \\ x_4 \\ \dfrac{\begin{pmatrix} -(2x_3x_4 + x_4^2 - x_3^2\cos x_2)\sin x_2 + 20\cos x_1 \\ -10\cos(x_1 + x_2)\cos x_2 \end{pmatrix}}{\cos^2 x_2 - 2} \\ \dfrac{\begin{pmatrix} (2x_3x_4 + x_4^2 + 2x_3x_4\cos x_2 + x_4^2\cos x_2 + 3x_3^2 \\ +2x_3^2\cos x_2 + 20(\cos(x_1 + x_2) - \cos x_1)(1 + \cos x_2) \\ -10\cos x_2\cos(x_1 + x_2) \end{pmatrix}}{\cos^2 x_2 - 2} \end{bmatrix}$ and $g(x) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ \dfrac{1}{2 - \cos^2 x_2} & \dfrac{-1 - \cos x_2}{2 - \cos^2 x_2} \\ \dfrac{-1 - \cos x_2}{2 - \cos^2 x_2} & \dfrac{3 + 2\cos x_2}{2 - \cos^2 x_2} \end{bmatrix}$.

The following simulation parameters were chosen for simulation, The initial system state vector was chosen as $\begin{bmatrix} \pi/6 & -\pi/6 & 0 & 0 \end{bmatrix}^T$. The cost function was selected as in (4) with $Q = I_{4\times4}$ and $R = I_{2\times2}$. The learning parameters were chosen as $\alpha_V = 0.025$, $\alpha_I = 0.055$, other design parameters were $g_M = 3$, $g_m = 1$, $\Gamma = 0.99$, $A = -10I$, and $P = I$ where I is the identify matrix. The basis function for approximating the value function is given by the expansion of $\phi(x) = [x_1^2; x_2^2; x_3^2; x_4^2; x_1x_2 \cdots; x_1^4; x_2^4; \cdots; x_1^3x_2; \cdots; x_1^2x_2x_3; \cdots; x_1x_2x_3x_4]$. The activation functions for the identifier $\sigma_I(\bullet) = diag\{\tanh(\bullet) \quad \tanh(\bullet)\}$. Number of hidden layer neurons for identifier and value function NN are selected as 25 and 39, respectively. All the NN weight estimates are initialized at random from a uniform distribution in the interval $(0, 1)$. The ultimate bound for the system state is chosen as 0.001.

The performance of the event sampled control system is shown in Figure 3. The system state is regulated close to zero as shown in Figure 3(a) along with the control input in Figure 3(b). The HJB equation error converges close to zero (shown in Figure 3 (c)) representing a near optimal solution is achieved with an event sampled implementation.

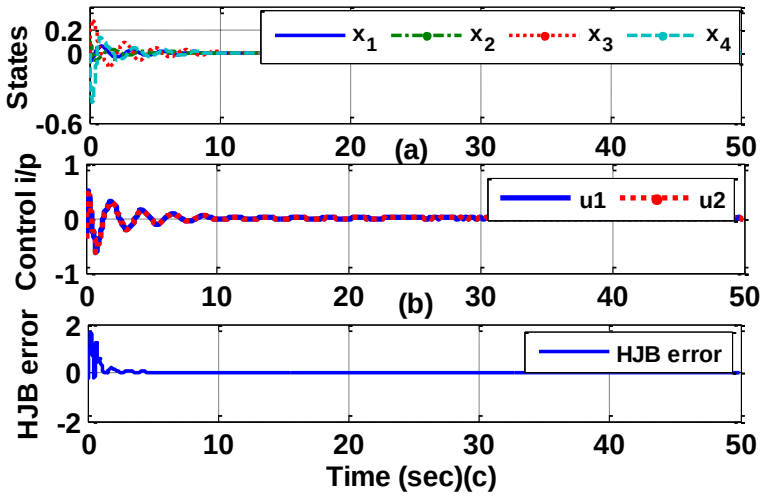

Figure 3. Convergence of (a) the system state; (b) control input (c) HJB error.

The evolution of the event sampling threshold along with the event sampling error is shown in Figure 4(a). Cumulative number of event sampled instants is shown in Figure 4 (b) and the inter-sample times in Figure 4(c). From the cumulative number of event sampled instants, it is evident that a fewer number of sampled instances occurred when compared to the traditional periodic sampled data system. The number of event sampled instants found to be 15783 for simulation duration of 50 sec with a sampling interval of 0.001 sec or 50,000 sampling instants for a traditional sampled data system.

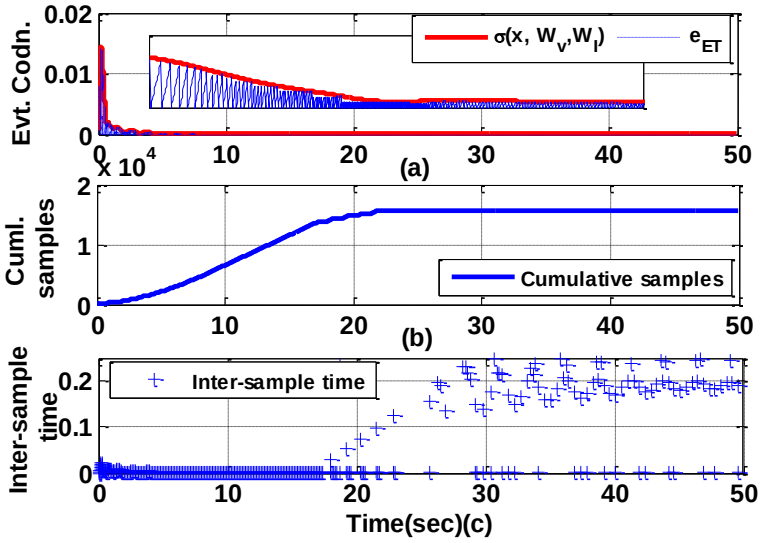

Figure 4. Evolution of (a) event sampling condition; (b) cumulative number of event sampled instants; (c) inter-sample times.

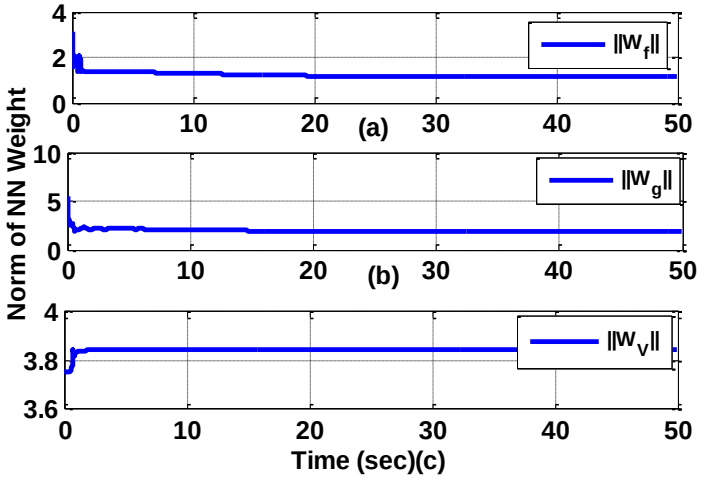

Figure 5. Convergence of the norm of the NN weight estimates.

It is further clear from the Figure 4 (b) that the event sampling condition generated a large number of sampled instants at the initial online NN learning phase. This is due to the large weight estimation error and makes the NN to learn the unknown system dynamics and the value function to achieve near optimality. Over time, as the NN

approximate the system dynamics and value function and the inter-sample times increased thereby reducing the number of sampled events. The convergence of all the NN weight estimates is shown in Figure 5 (a), (b) and (c).

6. CONCLUSIONS

In this paper, we proposed an event sampled near optimal control of an uncertain continuous-time system. A near optimal solution of the HJB equation is achieved with event sampled approximation of the value function and system dynamics. The NN weight tuning at the event sampled instants with adaptive event sampling condition is found to ensure convergence of the NN weight estimates to their respective target values. It was observed that the inter-sample times depend on the initial values of the NN weight estimates. Further, the inter-sample times found to increase with convergence of the parameters. The simulation results validated all the analytical design. The cost function considered in this paper only optimizes the control policy. The optimization of the event sampled instants will be an interesting problem and will be studied in the future.

7. REFERENCES

[1] F. L. Lewis and V. L. Syrmos, *Optimal Control*. Hoboken, NY: Wiley, 1995.

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

[3] Z. Chen and S. Jagannathan, "Generalized Hamilton-Jacobi-Bellman formulation based neural network control of affine nonlinear discrete time systems," *IEEE Transaction on Neural Networks*, vol. 19, no. 1, pp. 90-106, Jan. 2008.

[4] J. Si, A. G. Barto, W. B. Powell and D. Wunsch, *Handbook of Learning and Approximate Dynamics Programming*. New York: Wiley, 2004.

[5] D. Vrabie, K. Vamvoudakis, and F. Lewis, "Adaptive optimal controllers based on generalized policy iteration in a continuous-time framework," *in Proceedings of the17th IEEE Mediterranean Conference on Control and Automation*, Jun. 2009, pp. 1402-1409.

[6] Y. Jiang and Z. P. Jiang. Global adaptive dynamic programming for continuous-time nonlinear systems. *ArXiv:1401.0020 [math.DS]*, 2014. URL http://arxiv.org/pdf/1401.0020.pdf.

[7] D. V. Prokhorov and D. Wunsch, "Adaptive Critic Designs," *IEEE Transactions on Neural Networks*, vol 8, no. 5 , pp. 997–1007, Sep. 1997.

[8] D. Liu and Q. Wei, "Finite-approximation-error based optimal control approach for discrete-time nonlinear systems," *IEEE Transactionson Cybernetics*, vol. 43, no. 2, pp. 779-789, Apr. 2013.

[9] H. Zhang, Y. Luo, and D. Liu, "Neural-network-based near-optimal control for a class of discrete-time affine nonlinear systems with control constraints," *IEEE Transactions on Neural Network*, vol. 20, no. 9, pp. 1490–1503, Sep. 2009.

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

[11] Hao Xu, S. Jagannathan and F. L. Lewis, "Stochastic optimal control of unknown linear networked control system in the presence of random delays and packet losses," *Automatica*, vol. 48, no. 6, pp. 1017-1030, Jun. 2012.

[12] 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 Transaction on Neural Network and Learning Systems*, vol. 23, no. 7, pp. 1118-1129, Jul. 2012.

[13] 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 Transactions on Systems, Man, and Cybernetics—Part B*, vol. 38, no. 4, pp. 937-942, Aug. 2008.

[14] 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.

[15] R. Postoyan, A. Anta, D. Nesic, and P. Tabuada, "A unifying Lyapunov-based framework for event-triggered control of nonlinear systems," *in Proceedings of the 50th IEEE Conference on Decision and Control*, 2011, Orlando, FL, USA, pp. 2559–2564.

[16] S. Jagannathan, *Neural Network Control of Nonlinear Discrete-time Systems.* Boca Raton, FL: CRC Press, 2006.

[17] F. Lewis, S. Jagannathan and A. Yesildrek, *Neural Network Control of Robot Manipulators and Nonlinear Systems.* Bristol, PA*:* Taylor and Francis, 1998.

[18] H. K. Khalil, *Nonlinear Systems*. Upper Saddle river, NJ: Prentice Hall, 2002.

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

[20] R. Goebel, R. Sanfelice, and A. Teel, "Hybrid dynamical systems," *IEEE Control System Magazine*, vol. 29, no.2, pp. 28–93, Apr. 2009.

[21] T. Hayakawa and W. Haddad, "Stable neural hybrid adaptive control for nonlinear uncertain impulsive dynamical systems," *in Proceedings of the 44th IEEE Conference on Decision and Control, and the European Control Conference* , Seville, Spain, Dec. 2005, pp. 5510–5515.

[22] Q. Zhao, H. Xu, T. Dierks, and S. Jagannathan, "Finite-horizon neural network-based optimal control design for affine nonlinear continuous-time systems," in *Proceeding of the International Joint Conference on Neural Networks*, Dallas, TX, USA, Aug. 2013, pp. 1-6.

[23] A. Sahoo, H. Xu, and S. Jagannathan, "Event-triggered optimal control of nonlinear continuous-time systems in affine form by using neural networks," in *Proc. of the 53rd IEEE conference on Decision and Control*, Los Angeles, CA, USA, Dec. 2014, pp. 1277-1232.

APPENDIX

Proof of Lemma 1: Consider the term $\hat{\omega}\hat{\omega}^T$. By definition (30), it can be represented as

$$\hat{\omega}\hat{\omega}^T = \left(\nabla_x\phi(x)\hat{f}(x) - (1/2)\nabla_x\phi(x)\hat{D}(x)\nabla_x^T\phi(x)\hat{W}_V\right)\left(\nabla_x\phi(x)\hat{f}(x) - (1/2)\nabla_x\phi(x)\hat{D}(x)\nabla_x^T\phi(x)\hat{W}_V\right)^T.$$

Since $(a-b)^2 = (b-a)^2$, we can rewrite the above equation as

$$\hat{\omega}\hat{\omega}^T = \left((1/2)\nabla_x\phi(x)\hat{D}(x)\nabla_x^T\phi(x)\hat{W}_V - \nabla_x\phi(x)\hat{f}(x)\right)\left((1/2)\nabla_x\phi(x)\hat{D}(x)\nabla_x^T\phi(x)\hat{W}_V - \nabla_x\phi(x)\hat{f}(x)\right)^T. \quad \text{(A.1)}$$

Recalling the definition of the value function NN weight estimation error, we have $\hat{W}_V = W_V - \tilde{W}_V$. Substituting into (A.1) one can arrive at

$$\begin{aligned}\hat{\omega}\hat{\omega}^T = &\left((1/2)\nabla_x\phi(x)\hat{D}(x)\nabla_x^T\phi(x)W_V - \nabla_x\phi(x)\hat{f}(x) - (1/2)\nabla_x\phi(x)\hat{D}(x)\nabla_x^T\phi(x)\tilde{W}_V\right)\\ &\times\left((1/2)\nabla_x\phi(x)\hat{D}(x)\nabla_x^T\phi(x)W_V - \nabla_x\phi(x)\hat{f}(x) - (1/2)\nabla_x\phi(x)\hat{D}(x)\nabla_x^T\phi(x)\tilde{W}_V\right)^T.\end{aligned} \quad \text{(A.2)}$$

Moreover, by Young's inequality $2ab < (1/l)a^2 + lb^2$, $l > 0$. Using this relation it holds that

$$(a-b)^2 \geq (1-l)a^2 + (1-(1/l))b^2 \quad \text{(A.3)}$$

Using the above inequality, we have

$$\begin{aligned}\hat{\omega}\hat{\omega}^T \geq\ &(1-l)\left((1/2)\nabla_x\phi(x)\hat{D}(x)\nabla_x^T\phi(x)W_V - \nabla_x\phi(x)\hat{f}(x)\right)\\ &\times\left((1/2)\nabla_x\phi(x)\hat{D}(x)\nabla_x^T\phi(x)W_V - \nabla_x\phi(x)\hat{f}(x)\right)^T\\ &+(1-(1/l))(1/4)(\nabla_x\phi(x)\hat{D}(x)\nabla_x^T\phi(x)\tilde{W}_V)\times(\nabla_x\phi(x)\hat{D}(x)\nabla_x^T\phi(x)\tilde{W}_V)^T.\end{aligned}$$

By selecting $l = (N+1)/N$ where $N > 0$ is a positive integer, we have

$$\begin{aligned}\hat{\omega}\hat{\omega}^T \geq\ &-\frac{1}{N}\left((1/2)\nabla_x\phi(x)\hat{D}(x)\nabla_x^T\phi(x)W_V - \nabla_x\phi(x)\hat{f}(x)\right)\\ &\times\left((1/2)\nabla_x\phi(x)\hat{D}(x)\nabla_x^T\phi(x)W_V - \nabla_x\phi(x)\hat{f}(x)\right)^T\\ &+\frac{1}{4(N+1)}(\nabla_x\phi(x)\hat{D}(x)\nabla_x^T\phi(x)\tilde{W}_V)(\nabla_x\phi(x)\hat{D}(x)\nabla_x^T\phi(x)\tilde{W}_V)^T.\end{aligned} \quad \text{(A.4)}$$

Pre and post multiplying $\tilde{W}_V^T$ and $\tilde{W}_V$ both side of (A.4), respectively, and changing the sign one can reach at (46) of Lemma 1. ■

Proof of Lemma 2: Since we assume that a minimum inter-sample time exists between two consecutive sampling instants, the jumps occur at distinct time instants. Thus, the boundedness of the NN identifier weight estimation errors is proven by considering both the flow duration and jump instants as in [19].

Case 1: During the flow for $t_k < t \le t_{k+1},\ k = 1, 2, \cdots$.

Consider the Lyapunov function candidate given as

$$L_{\tilde{W}} = \beta_{I,1} L_I + L_V + \beta_{I,2} L_{I,2}, \tag{A.5}$$

where $L_I = tr\{\tilde{W}_I^T \tilde{W}_I\}$, $L_V = \tilde{W}_V^T \tilde{W}_V$, and $L_{I,2} = tr\{\tilde{W}_I^T \tilde{W}_I\}^2$ with $\beta_{I,1} = 2\varpi_3 / \alpha_I (1 - 3\alpha_I)$ and $\beta_{I,2} = 2\varpi_2 / \varpi_4$ with ϖ_1, ϖ_2, ϖ_3 and ϖ_4 are positive constants defined during the proof.

The first derivative of the Lyapunov function candidate (A.5) by using the weight estimation error dynamics (22) and (31) becomes

$$\dot{L}_{\tilde{W}} = 2\beta_{I,1} tr\{\tilde{W}_I^T \dot{\tilde{W}}_I\} + 2\dot{\tilde{W}}_V^T \tilde{W}_V + 4\beta_{I,2} tr\{\tilde{W}_I^T \tilde{W}_I\} tr\{\dot{\tilde{W}}_I^T \tilde{W}_I\} = 0. \tag{A.6}$$

From (A.6), the first derivative $\dot{L}_{\tilde{W}}$ is zero, which implies the NN weight estimation errors remain constant during the flow for $t_k < t \le t_{k+1},\ \forall k = 1, 2, \cdots$. Since, the initial NN weights $\hat{W}_I(0)$ and $\hat{W}_V(0)$, and the target NN weights W_I and W_V, are bounded, the initial weight estimation errors $\tilde{W}_I(0)$ and $\tilde{W}_V(0)$ are bounded. Therefore, to prove the boundedness of $\tilde{W}_I$ and $\tilde{W}_V$ for all time we only need to show that $\tilde{W}_I$ and $\tilde{W}_V$ are bounded during the jump instants.

Case 2: At the jumps ($t = t_k$, $\forall k = 1, 2 \cdots$)

Consider (A.5) in Case I as a discrete Lyapunov function candidate. We will consider each term in (A.5) individually to evaluate the first difference at jump instants and combine them to reach the overall first difference to prove locally UB.

Consider the first term of the Lyapunov function candidate (A.5). The first difference is given by

$$\Delta L_I = tr\left\{\tilde{W}_I^{T+}\tilde{W}_I^{+}\right\} - tr\left\{\tilde{W}_I^{T}\tilde{W}_I\right\}. \tag{A.7}$$

Recalling the dynamics of the weight estimation error (21), the first difference ΔL_I of the Lyapunov function becomes

$$\Delta L_I = -2\alpha_I tr\left\{\tilde{W}_I^T \sigma_I(x)\chi_I \bar{u}e_I^T\right\} - 2\alpha_I^2 tr\left\{(\sigma_I(x)\chi_I \bar{u}e_I^T)^T \hat{W}_I\right\} + 2\alpha_I tr\left\{\tilde{W}_I^T \hat{W}_I\right\}$$
$$+\alpha_I^2\chi_I^2 tr\left\{(\sigma_I(x)\bar{u}e_I^T)^T(\sigma_I(x)\bar{u}e_I^T)\right\} + \alpha_I^2 tr\left\{\hat{W}_I^T\hat{W}_I\right\},$$

where $\chi_I = 1/(c + \|e_I\|^2)$. By replacing $\hat{W}_I = W_I - \tilde{W}_I$ and using the inequality $tr(AB) \le (1/2)\|A\|^2 + (1/2)\|B\|^2$, the first difference is upper bound by

$$\Delta L_I \le 2\alpha_I \left\|\tilde{W}_I\right\|^2 + \frac{\alpha_I}{2}\sigma_{I,M}^2\chi_I^2\|\bar{u}\|^2\|e_I\|^2 + \alpha_I^2\left\|\tilde{W}_I\right\|^2 + \alpha_I^2\sigma_{I,M}^2\chi_I^2\|\bar{u}\|^2\|e_I\|^2 + \alpha_I^2 W_{I,M}^2$$
$$+\alpha_I^2\sigma_{I,M}^2\chi_I^2\|\bar{u}\|^2\|e_I\|^2 - 2\alpha_I\left\|\tilde{W}_I\right\|^2 + \frac{\alpha_I}{2}\left\|\tilde{W}_I\right\|^2 + 2\alpha_I W_{I,M}^2 + \alpha_I^2\chi_I^2\sigma_{I,M}^2\|\bar{u}\|^2\|e_I\|^2$$
$$+2\alpha_I^2 W_{I,M}^2 + 2\alpha_I^2\left\|\tilde{W}_I\right\|^2.$$

Observe that $\chi_I\|\bar{u}\|\|e_I\| = \|\bar{u}\|\|e_I\|/(c + \|\bar{u}\|^2\|e_I\|^2) < 1$ and combining similar terms we arrive at

$$\Delta L_I \le -\alpha_I(1 - 3\alpha_I)\left\|\tilde{W}_I\right\|^2 + B_{W,I}, \tag{A.8}$$

where $B_{W,I} = (2\alpha_I + 3\alpha_I^2)W_{I,M}^2 + (\alpha_I + 3\alpha_I^2)\sigma_{I,M}^2$.

Next, considering the second term L_V , the first difference $\Delta L_V = (\tilde{W}_V^+)^T \tilde{W}_V^+ - \tilde{W}_V^T \tilde{W}_V$, along the dynamics of the weight update law (32) can be expressed as

$$\Delta L_V = \frac{2\alpha_V \tilde{W}_V^T \hat{\omega} \hat{H}^{+T}(x,u)}{(1+\hat{\omega}^T\hat{\omega})^2} + \left[\frac{\alpha_V \hat{\omega} \hat{H}^{+T}(x,\hat{W}_V)}{(1+\hat{\omega}^T\hat{\omega})^2}\right]^T \left[\frac{\alpha_V \hat{\omega} \hat{H}^{+T}(x,\hat{W}_V)}{(1+\hat{\omega}^T\hat{\omega})^2}\right].$$

Substituting $\hat{H}^+(x,\hat{W}_V)$ from (35) and applying Young's inequality $2a^Tb \le (1/p)a^Ta + pb^Tb$, along with the relation $\hat{\omega}^T\hat{\omega}/(1+\hat{\omega}^T\hat{\omega})^2 \le 1$, the first difference is upper bounded by

$$\begin{aligned}
\Delta L_V \le & -\alpha_V((1/8)-5\alpha_V)\frac{\tilde{W}_V^T\hat{\omega}\hat{\omega}^T\tilde{W}_V}{(1+\hat{\omega}^T\hat{\omega})^2} - \frac{\alpha_V}{2}\frac{\tilde{W}_V^T\hat{\omega}\hat{\omega}^T\tilde{W}_V}{(1+\hat{\omega}^T\hat{\omega})^2} \\
& + \frac{\alpha_V(8+5\alpha_V)W_V^T\nabla_x\phi(x)\tilde{f}(x))\tilde{f}^T(x)\nabla_x^T\phi(x)W_V}{(1+\hat{\omega}^T\hat{\omega})^2} \\
& + \frac{\alpha_V(1+5\alpha_V)}{(1+\hat{\omega}^T\hat{\omega})^2}((1/4)\tilde{W}_V^T\nabla_x\phi(x)\hat{D}(x)\nabla_x^T\phi(x)\tilde{W}_V)((1/4)\tilde{W}_V^T\nabla_x\phi(x)\hat{D}(x)\nabla_x^T\phi(x)\tilde{W}_V)^T \qquad \text{(A.9)} \\
& + \frac{\alpha_V(8+5\alpha_V)}{(1+\hat{\omega}^T\hat{\omega})^2}((1/4)W_V^T\nabla_x\phi(x)\tilde{D}(x)\nabla_x\phi^T(x)W_V)((1/4)W_V^T\nabla_x\phi(x)\tilde{D}(x)\nabla_x\phi^T(x)W_V)^T \\
& + \left(\alpha_V(8+5\alpha_V)\varepsilon_H\varepsilon_H^T/(1+\hat{\omega}^T\hat{\omega})^2\right).
\end{aligned}$$

Recall the Lemma 1 and multiply $\alpha_V/2$ both side of (46). Substituting into (A.9), and applying Frobenius norm and Young's inequality $2ab \le a^2 + b^2$, reveals that

$$\begin{aligned}
\Delta L_V \le & -\alpha_V((1/8)-5\alpha_V)\frac{\tilde{W}_V^T\hat{\omega}\hat{\omega}^T\tilde{W}_V}{(1+\hat{\omega}^T\hat{\omega})^2} - \frac{\alpha_V}{8\left(1+\hat{\omega}^T\hat{\omega}\right)^2}\left(\frac{1}{(N+1)} - \frac{(1+5\alpha_V)}{2}\right) \\
& \times\left\|\tilde{W}_V^T\nabla_x\phi(x)\hat{D}(x)\nabla_x^T\phi(x)\tilde{W}_V\right\|^2 + \frac{D_{\min}\alpha_V^2}{N\left(1+\hat{\omega}^T\hat{\omega}\right)^2}\left\|\tilde{W}_V\right\|^4 + \frac{1}{D_{\min}N\left(1+\hat{\omega}^T\hat{\omega}\right)^2} \\
& \times\left\|\left((1/2)\nabla_x\phi(x)\hat{D}(x)\nabla_x^T\phi(x)W_V - \nabla_x\phi(x)\hat{f}(x)\right)\right\|^4 + \frac{\alpha_V(8+5\alpha_V)}{(1+\hat{\omega}^T\hat{\omega})^2}
\end{aligned}$$

$$\times\left\|W_V^T\nabla_x\phi(x)\,\tilde{f}(x))\,\tilde{f}^T(x)\nabla_x^T\phi(x)W_V\right\|+\frac{\alpha_V(8+5\alpha_V)}{16(1+\hat{\omega}^T\hat{\omega})^2}\left\|W_V^T\nabla_x\phi(x)\tilde{D}(x)\nabla_x\phi^T(x)W_V\right\|^2$$

$$+\frac{\alpha_V(8+5\alpha_V)\left\|\varepsilon_H\right\|^2}{(1+\hat{\omega}^T\hat{\omega})^2}.$$

where $\bar{D}_{\min}$ is a positive constant and satisfies the inequality $0<\bar{D}_{\min}\le\left\|\nabla_x\phi(x)\hat{D}(x)\nabla_x^T\phi(x)\right\|$. Note that the gradient of the activation function, $\nabla_x\phi(x)$, of the value function neural network satisfies PE condition. Further, from (20) the function $\hat{D}(x)=\hat{g}(x)R^{-1}\hat{g}^T(x)$ is lower bounded. Therefore, it holds that $D_{\min}$ is a positive constant. By using this relation and applying Cauchy inequality $(a+b)^4\le 4a^4+4b^4$ with one can reach at

$$\Delta L_V\le-\alpha_V((1/8)-5\alpha_V)\frac{\tilde{W}_V^T\hat{\omega}\hat{\omega}^T\tilde{W}_V}{(1+\hat{\omega}^T\hat{\omega})^2}-\frac{\alpha_V D_{\min}}{8\left(1+\hat{\omega}^T\hat{\omega}\right)^2}\left(\frac{1}{(N+1)}-\frac{(1+5\alpha_V)}{2}-\frac{8\alpha_V}{N}\right)\left\|\tilde{W}_V\right\|^4$$

$$+\frac{1}{4D_{\min}N}\phi_M^{'8}W_{V,M}^4\left\|\hat{D}(x)\right\|^4+\frac{4}{D_{\min}N}\phi_M^{'4}\left\|\hat{f}(x)\right\|^4+\alpha_V(8+5\alpha_V)\phi_M^{'2}W_{V,M}^2\left\|\tilde{f}(x)\right\|^2$$

$$+\alpha_V(8+5\alpha_V)\phi_M^{'4}W_{V,M}^4\left\|\tilde{D}(x)\right\|^2+\alpha_V(8+5\alpha_V)\varepsilon_{H,M}^2.$$

Recalling the dynamics of the identifier we have the following conclusions:

$$\left\|\tilde{f}(x)\right\|\le\left\|[\tilde{f}(x)\;\;\tilde{g}(x)]\right\|=\left\|\tilde{W}_I^T\sigma_I(x)+\varepsilon_I(x)\right\|,$$

$$\begin{aligned}\tilde{D}(x)&=D(x)-\hat{D}(x)=g(x)R^{-1}g^T(x)-\hat{g}(x)R^{-1}\hat{g}^T(x)\\&=\tilde{g}(x)R^{-1}g^T(x)+g(x)R^{-1}\tilde{g}^T(x)-\tilde{g}(x)R^{-1}\tilde{g}^T(x)\\&\le 2\lambda_M(R^{-1})g_M\left\|\tilde{W}_I^T\sigma_I(x)+\varepsilon_I(x)\right\|,\end{aligned}$$

$$\left\|\hat{f}(x)\right\|\le\left\|[\hat{f}(x)\;\;\hat{g}(x)]\right\|=\left\|\hat{W}_I^T\sigma_I(x)\right\|,\text{ and}$$

$$\left\|\hat{D}(x)\right\|\le D_M+\left\|\tilde{D}(x)\right\|\le D_M+2\lambda_M\left(R^{-1}\right)g_M\left\|\tilde{W}_I^T\sigma_I(x)+\varepsilon_I(x)\right\|.$$

Substituting the above facts, the first difference is upper bounded by

$$\Delta L_V \le -\alpha_V((1/8)-5\alpha_V)\frac{\tilde{W}_V^T\hat{\omega}\hat{\omega}^T\tilde{W}_V}{(1+\hat{\omega}^T\hat{\omega})^2}-\frac{\alpha_V D_{\min}}{8\left(1+\hat{\omega}^T\hat{\omega}\right)^2}\left(\frac{1-N}{2(N+1)}-\frac{(16+5N)\alpha_V}{2N}\right)\left\|\tilde{W}_V\right\|^4$$

$$+\frac{2}{D_{\min}N}\left(\phi_M'^8W_{V,M}^4\lambda_M^4\left(R^{-1}\right)g_M^4\sigma_{I,M}^4+8\phi_M'^4\sigma_{I,M}^4\right)\left\|\tilde{W}_I\right\|^4+\left(2\alpha_V(8+5\alpha_V)\phi_M'^2W_{V,M}^2\sigma_{I,M}^2\right. \quad \text{(A.10)}$$

$$\left.+8\alpha_V(8+5\alpha_V)\phi_M'^4W_{V,M}^4\lambda_M^2\left(R^{-1}\right)g_M^2\sigma_{I,M}^2\right)\left\|\tilde{W}_I\right\|^2+\varepsilon_{V,2}$$

where $\varepsilon_{V,2}=8\alpha_V(8+5\alpha_V)\phi_M'^4W_{V,M}^4\lambda_M^2(R^{-1})g_M^2\varepsilon_{I,M}^2+2\alpha_V(8+5\alpha_V)\phi_M'^2W_{V,M}^2\varepsilon_{I,M}^2$

$$+\frac{16}{D_{\min}N}\phi_M'^4\varepsilon_{I,M}^4+\frac{2}{D_{\min}N}\phi_M'^8W_{V,M}^4\lambda_M^4(R^{-1})g_M^4\varepsilon_{I,M}^4+\frac{1}{D_{\min}N}\phi_M'^8W_{V,M}^4D_M^4+\alpha_V(8+5\alpha_V)\varepsilon_{H,M}^2$$

and $\frac{1-N}{2(N+1)}-\frac{(16+5N)\alpha_V}{2N}>0$ by selecting $0<\alpha_V<\min\left\{\frac{1}{40},\frac{2N(1-N)}{2(N+1)(16+N)}\right\}$ with $0<N<1$.

Considering the third term $L_{I,2}=tr\{\tilde{W}_I^T\tilde{W}_I\}^2$, the first difference can be written as

$$\Delta L_{I,2}=tr\{\tilde{W}_I^{T+}\tilde{W}_I^{+}\}^2-tr\{\tilde{W}_I^T\tilde{W}_I\}^2$$

$$=\Delta L_I\left(\Delta L_I+2tr\{\tilde{W}_I^T\tilde{W}_I\}\right)\le\left(\Delta L_I\right)^2+2\Delta L_I\left\|\tilde{W}_I\right\|^2. \quad \text{(A.11)}$$

Substituting the first difference (A.8), the $\Delta L_{I,2}$ can be expressed as

$$\Delta L_{I,2}\le-[2\alpha_I(1-3\alpha_I)(1-\alpha_I(1-3\alpha_I))+1]\left\|\tilde{W}_I\right\|^4+3B_{W,I}^2. \quad \text{(A.12)}$$

At the final step, combining all the individual first differences, (A.8), (A.10), (A.10), and (A.12), the overall first difference of the Lyapunov function (A.5) becomes

$$\Delta L_{\tilde{W}}\le-\alpha_V((1/8)-5\alpha_V)\frac{\tilde{W}_V^T\hat{\omega}\hat{\omega}^T\tilde{W}_V}{(1+\hat{\omega}^T\hat{\omega})^2}-\varpi_1\left\|\tilde{W}_V\right\|^4+\varpi_2\left\|\tilde{W}_I\right\|^4+\varpi_3\left\|\tilde{W}_I\right\|^2$$

$$-\beta_{I,1}\alpha_I(1-3\alpha_I)\left\|\tilde{W}_I\right\|^2-\beta_{I,2}\varpi_4\left\|\tilde{W}_I\right\|^4+\varepsilon_{\tilde{W}},$$

where $$\varpi_1=\frac{\alpha_V D_{\min}}{8\left(1+\hat{\omega}^T\hat{\omega}\right)^2}\left(\frac{1-N}{2(N+1)}-\frac{(16+5N)\alpha_V}{2N}\right),$$

$$\varpi_2 = \frac{2\left(\phi_M'^8 W_{V,M}^4 \lambda_M^4 \left(R^{-1}\right) g_M^4 \sigma_{I,M}^4 + 8\phi_M'^4 \sigma_{I,M}^4\right)}{D_{\min} N}$$

$$\varpi_3 = \left(2\alpha_V(8+5\alpha_V)\phi_M'^2 W_{V,M}^2 \sigma_{I,M}^2 + 8\alpha_V(8+5\alpha_V)\phi_M'^4 W_{V,M}^4 \lambda_M^2 \left(R^{-1}\right) g_M^2 \sigma_{I,M}^2\right)$$

$$\varpi_4 = [2\alpha_I(1-3\alpha_I)(1-\alpha_I(1-3\alpha_I))+1]$$

and $\varepsilon_{\tilde{W}} = 3\beta_{I,2} B_{W,I}^2 + \beta_{I,1} B_{W,I} + \varepsilon_{V,2}$ are the new variables introduced for simplicity.

Recalling the definition of $\beta_{I,1} = 2\varpi_3/\alpha_I(1-3\alpha_I)$ and $\beta_{I,2} = 2\varpi_2/\varpi_4$, the first difference is upper bounded by

$$\Delta L_{\tilde{W}} \le -\varpi_1 \left\|\tilde{W}_V\right\|^4 - \varpi_2 \left\|\tilde{W}_I\right\|^4 - \varpi_3 \left\|\tilde{W}_I\right\|^2 + \varepsilon_{\tilde{W}}. \tag{A.13}$$

From (A.13), the first difference $\Delta L_{\tilde{W}} < 0$ as long as

$$\left\|\tilde{W}_I\right\| > \max\{\sqrt{\varepsilon_{\tilde{W}}/\alpha_I(1-3\alpha_I)}, \sqrt[4]{\varepsilon_{\tilde{W}}/\varpi_2}\} = B_{\tilde{W}_I} \text{ or}$$

$$\left\|\tilde{W}_V\right\| > \sqrt[4]{\varepsilon_{\tilde{W}}/\varpi_1} = B_{\tilde{W}_V}.$$

Hence, the NN weight estimation errors $\tilde{W}_I$ and $\tilde{W}_V$ are bounded at the jump instants.

From Case 1 and Case 2, the NN weight estimation errors $\tilde{W}_I$ and $\tilde{W}_V$ remains constant during the flow and are locally ultimately bounded during jump. Therefore, $\tilde{W}_I$ and $\tilde{W}_V$ are locally ultimately bounded for all $t_k > \bar{T}$ or for $t > T$ for $T > \bar{T}$ with an ultimate bound $B_{\tilde{W}}^{ub} = \max\{B_{\tilde{W}_I}, B_{\tilde{W}_V}\}$. ■

Proof of Theorem 1: For proving the Theorem 1 we need to show the nonlinear impulsive dynamical system is locally ultimately bounded both during flow period and jump instants as discussed in [20].

Case 1: *Flow period (i.e.* $t_k < t \le t_{k+1},\ \forall k = 1, 2, \cdots$*)*

Consider the Lyapunov function candidate given as

$$L_{cl} = L_x + L_{e_I} + L_{\bar{x}} + L_{\tilde{W}}, \tag{A.14}$$

where $L_x = V(x)$, $L_{e_I} = e_I^T P e_I$, $L_{\bar{x}} = \bar{x}^T \bar{x}$, $L_I = tr\{\tilde{W}_I^T \tilde{W}_I\}$, and $L_{\tilde{W}}$ is defined as in Lemma 2.

Considering the first term the first derivative along the closed loop system dynamics

$$\dot{L}_x = \nabla_x V(x)\dot{x} = \nabla_x V(x)(f(x) + g(x)u).$$

Substituting the closed-loop dynamics from (37), the first derivative leads to

$$\begin{aligned}\dot{L}_x &= \nabla_x V(x)(f(x) + g(x)u^*) + \nabla_x V(x)[(1/2)g(x)R^{-1}\tilde{g}(x)\nabla_x^T\phi(x)W_V \\ &+(1/2)g(x)R^{-1}\hat{g}(x)\nabla_x^T\phi(x)\tilde{W}_V + (1/2)g(x)R^{-1}\hat{g}(x)(\nabla_x^T\phi(x) - \nabla_x^T\phi(\bar{x}))\hat{W}_V + \varepsilon_u].\end{aligned}$$

Recall the definition of the optimal value function, then it holds that $\nabla_x V(x)(f(x) + g(x)u^*) \le -Q(x)$. Substituting the above inequality and using the Lipschitz continuity of the gradient of the value function activation function one can arrive at

$$\begin{aligned}\dot{L}_x &\le -Q(x) + (1/2)g_M\lambda_{\max}(R^{-1})\|\nabla_x V(x)\|\|\hat{g}(x)\|C_{\nabla\phi}\|e_{ET}\| + (1/2)\|\nabla_x V(x)\| \\ &\times\|g(x)\|\|R^{-1}\|\|\tilde{g}(x)\|\|\nabla_x^T\phi(x)W_V\| + (1/2)\|\nabla_x V(x)\|\|g(x)\|\|R^{-1}\|\|\hat{g}(x)\|\|\nabla_x^T\phi(x)\tilde{W}\| \\ &+\|\nabla_x V(x)\varepsilon_u\|.\end{aligned}$$

By Assumption 1 $\|g(x)\| \le g_M$ and using the NN approximation, $\|\tilde{g}(x)\| \le \|\tilde{W}_I^T\sigma_I(x) + \varepsilon_I(x)\|$, $\|\hat{g}(x)\| \le \|\hat{W}_I^T\sigma_I(x)\|$, and $\|\nabla_x V(x)\| = \|\nabla_x^T\phi(x)W + \nabla\varepsilon_V(x)\|$. Using the above facts and applying Young's inequality $2ab \le a^2 + b^2$, the first derivative satisfies

$$\dot{L}_x \le -Q(x) + (1/2) g_M \lambda_{\max}(R^{-1}) \sigma_{I,M} \phi_M^{'} \left\| \hat{W}_V \right\| \left\| \hat{W}_I \right\| C_{\nabla\phi} \left\| e_{ET} \right\| + (1/4) C_{\nabla\phi} g_M \lambda_{\max}(R^{-1}) \times \sigma_{I,M} \phi_M^{'} \left\| \hat{W}_I \right\| \left\| e_{ET} \right\|^2 + B_{F1}, \tag{A.15}$$

where

$$B_{F1} = (1/4) C_{\nabla\phi} g_M \lambda_{\max}(R^{-1}) \sigma_{I,M} \phi_M^{'} \left\| \tilde{W}_I \right\| \left\| \tilde{W}_V \right\|^2 + (1/4) C_{\nabla\phi} g_M \lambda_{\max}(R^{-1}) \sigma_{I,M} \phi_M^{'} W_{I,M} \left\| \tilde{W}_V \right\|^2$$
$$+ (1/2) g_M \lambda_{\max}(R^{-1}) (\phi_M^{'} W_{V,M} + \varepsilon_{V,M}^{'}) (\sigma_{I,M} \left\| \tilde{W}_I \right\| + \varepsilon_{I,M}) \phi_M^{'} W_{V,M}$$
$$+ (1/2) g_M \lambda_{\max}(R^{-1}) (\phi_M^{'} W_{V,M} + \varepsilon_{V,M}^{'}) \sigma_{I,M} \phi_M^{'} W_{I,M} \left\| \tilde{W}_V \right\|$$
$$+ (1/2) g_M \lambda_{\max}(R^{-1}) (\phi_M^{'} W_{V,M} + \varepsilon_{V,M}^{'}) \sigma_{I,M} \phi_M^{'} \left\| \tilde{W}_I \right\| \left\| \tilde{W}_V \right\| + (\phi_M^{'} W_{V,M} + \varepsilon_{V,M}^{'}) \varepsilon_{u,M}.$$

Considering the second term $L_{e_I} = e_I^T P e_I$, the first derivative along the identifier error dynamics (17) can be expressed as

$$\begin{aligned} \dot{L}_{e_I} &= \dot{e}_I^T P e_I + e_I^T P \dot{e}_I \\ &= e_I^T (A^T P + PA) e_I - 2 e_{ET}^T A P e_I + 2 (\tilde{W}_I^T \sigma_I(\breve{x}) \bar{u} + \hat{W}_I^T \tilde{\sigma}_I(x, \breve{x}) \bar{u} + \varepsilon_I)^T P e_I \\ &= -e_I^T \Pi e_I - 2 e_{ET}^T A P e_I + 2 (\tilde{W}_I^T \sigma_I(\breve{x}) \bar{u} + \hat{W}_I^T \tilde{\sigma}_I(x, \breve{x}) \bar{u} + \varepsilon_I)^T P e_I. \end{aligned}$$

By triangle inequality, the first derivative leads to

$$\dot{L}_{e_I} \le -\Pi_{\min} \left\| e_I \right\|^2 + 2 \left\| e_{ET} \right\| \left\| AP \right\| \left\| e_I \right\| + \left\| 2 (\tilde{W}_I^T \sigma_I(\breve{x}) \bar{u})^T P e_I \right\| + \left\| 2 (\hat{W}_I^T \tilde{\sigma}_I(x, \breve{x}) \bar{u})^T P e_I \right\| + \left\| 2 (\varepsilon_I)^T P e_I \right\|,$$

where $A^T P + PA = -\Pi$ is the Lyapunov equation and $\Pi_{\min} = \lambda_{\min}(\Pi)$ with $\lambda_{\min}(\bullet)$ is the minimum eigenvalue.

Applying Young's inequality $2ab \le (1/q) a^2 + q b^2$ and recalling the Lipschitz continuity of the identifier activation function, one can arrive at

$$\dot{L}_{e_I} \le -\frac{1}{2} \Pi_{\min} \left\| e_I \right\|^2 + (8/\Pi_{\min}) \left(C_{\sigma_I} \left\| \hat{W}_I \right\|^2 \left\| \bar{u} \right\|^2 \left\| P \right\|^2 + \left\| AP \right\|^2 \right) \left\| e_{ET} \right\|^2$$
$$+ (8/\Pi_{\min}) \left\| \tilde{W}_I^T \sigma_I(\breve{x}) \bar{u} \right\|^2 \left\| P \right\|^2 + (8/\Pi_{\min}) \left\| P \right\|^2 \varepsilon_{I,M}^2. \tag{A.16}$$

Next, the first derivatives of the third term can be expressed as

$$\dot{L}_{\breve{x}} = \dot{\breve{x}}^T \breve{x} + \breve{x}^T \dot{\breve{x}} = 0. \tag{A.17}$$

Considering the last term $L_{\tilde{W}}$, the derivative remains same as in (A.6) of Lemma 2 and given as

$$\dot{L}_{\tilde{W}} = 2\beta_{I,1} tr\{\tilde{W}_I^T \dot{\tilde{W}}_I\} + 2\dot{\tilde{W}}_V^T \tilde{W}_V + 4\beta_{I,2} tr\{\tilde{W}_I^T \tilde{W}_I\} tr\{\dot{\tilde{W}}_I^T \tilde{W}_I\} = 0. \tag{A.18}$$

Finally combining all individual first derivatives (A.15), (A.16), (A.17), and (A.18) the first derivative of the Lyapunov function becomes

$$\begin{aligned} \dot{L}_{cl} \le & -Q(x) - (1/2)\Pi_{\min} \left\|e_I\right\|^2 + (1/2) g_M \lambda_{\max}(R^{-1}) \sigma_{I,M} \phi_M^{'} \left\|\hat{W}_V\right\| \left\|\hat{W}_I\right\| C_{\nabla\phi} \left\|e_{ET}\right\| \\ & + (1/4)\left(C_{\nabla\phi} g_M \lambda_{\max}(R^{-1}) \sigma_{I,M} \phi_M^{'} \left\|\hat{W}_I\right\| + (32/\Pi_{\min})(C_{\sigma_I} \left\|\hat{W}_I\right\|^2 \|\bar{u}\|^2 \|P\|^2 + \|AP\|^2) \right) \\ & \times \left\|e_{ET}\right\|^2 + B_{cl,1k}. \end{aligned} \tag{A.19}$$

where $B_{cl,1k} = (8/\Pi_{\min}) \left\|\tilde{W}_I^T \sigma_I(\bar{x}) \bar{u}\right\|^2 \|P\|^2 + (8/\Pi_{\min}) \|P\|^2 \varepsilon_{I,M}^2 + B_{F1}$.

Recalling the event sampling condition (44) and substituting in (A.19). The first derivative leads to

$$\dot{L}_{cl} \le -(1-\Gamma) Q(x) - (1/2)\Pi_{\min} \left\|e_I\right\|^2 + B_{cl,1k}, \tag{A.20}$$

Observe that $B_{cl,1k}$ is a piece wise constant function since the NN weights are not updated during the flow duration and $\left\|\tilde{W}_I\right\|$ and $\left\|\tilde{W}_V\right\|$ are constant during each k^{th} flow duration.

From (A.20) it is clear that the first derivative of the Lyapunov function $\dot{L}_{cl} < 0$ as long as

$$Q(x) > B_{cl,1k} / (1-\Gamma) \equiv B_k^Q \text{ or}$$

$$\left\|e_I\right\| > \sqrt{B_{cl,1k} / (1/2)\Pi_{\min}} \equiv B_k^{e_I}.$$

Since $B_{cl,1k}$ is a piece wise constant function the bounds B_k^Q and $B_k^{e_I}$ are also piece wise constant during functions and remains constant during the flow period $t_k < t \le t_{k+1}$ for each $k = 1, 2, \cdots$.

Recalling the fact, $Q(x)$ is a positive definite quadratic function of x, with $Q(0) = 0$ and $Q(x) \to \infty$ as $x \to \infty$, the closed-loop system state, x, is bounded during flow period with a bound $B_k^x = Q^{-1}\left(B_k^Q\right)$. Further, as the NN weights for the value function and the identifier are not updated during the flow period the NN weight estimation errors $\tilde{W}_V$ and $\tilde{W}_I$, respectively, remain constant and bounded. Therefore, the states of the impulsive dynamical system ζ remains bounded during the flow period.

Furthermore, from Lemma 2, the weight estimation errors $\tilde{W}_V$ and $\tilde{W}_I$ converge to the ultimate bound $B_{\tilde{W}}^{ub}$ for all $t_k > \bar{T}$. Therefore, the system state x and the last held state $\breve{x}$ converge to the ultimate bound given by $B_{ub}^x = Q^{-1}\left(B_{ub}^Q\right)$ where $B_{ub}^Q = B_{cl,1}^{ub} / (1 - \Gamma)$ where $B_{cl,1}^{ub}$ computed from (A.20) by replacing $\tilde{W}_V$ and $\tilde{W}_I$ with their ultimate bounds $B_{\tilde{W}}^{ub}$ in the expression for $B_{cl,1k}$. Similarly the ultimate bound for the identification error is given by $B_{ub}^{e_I} = \sqrt{B_{cl}^{ub} / (1/2)\Pi_{\min}}$. Therefore, the closed-loop nonlinear impulsive dynamical system state ξ locally UB for all $t_k > \bar{T}$ with an ultimate bound $\mu_f = \max\{B_{ub}^x, B_{ub}^{e_I}, B_{\tilde{W}}^{ub}\}$. Next, it remains to prove that the closed-loop signals are bounded during jump.

Case 2: *When an event is triggered (i.e. $t = \tau_k$)*

Consider the discrete Lyapunov function candidate given by

$$L_{cl} = L_x + L_{e_I} + L_{\breve{x}} + L_{\tilde{W}}, \tag{A.21}$$

where $L_x = V(x)$, $L_{e_I} = e_I^T P e_I$, $L_{\breve{x}} = \breve{x}^T \breve{x}$, and $L_{\tilde{W}} = \beta_{I,1} L_I + L_V + \beta_{I,2} L_{I,2}$ as in Lemma 2. The first difference of Lyapunov function candidate can be represented as

$$\Delta L_{cl} = \Delta L_x + \Delta L_{e_I} + \Delta L_{\breve{x}} + \Delta L_{\tilde{W}} \tag{A.22}$$

Consider the first term of the first difference (A.22). Along the jump dynamics (40), the first difference

$$\Delta L_x = V(x^+) - V(x) = 0. \tag{A.23}$$

Similarly, the first difference of the second term along the identification error dynamics (41) becomes

$$\Delta L_{e_I} = e_I^{+T} P e_I^+ - e_I^T P e_I = e_I^T P e_I - e_I^T P e_I = 0. \tag{A.24}$$

Consider the second term $L_{\breve{x}} = \breve{x}^T \breve{x},$, the first difference can be expressed as

$$\Delta L_{\breve{x}} = \breve{x}^{+T} \breve{x}^+ - \breve{x}^T \breve{x} = x^T x - \breve{x}^T \breve{x} \le -\|\breve{x}\|^2 + B_k^x. \tag{A.25}$$

where $B_k^x = Q^{-1}\left(B_k^Q\right)$ is the bound during each flow interval defined earlier in Case I.

Next, the first difference of the rest of the terms can be written from using Lemma 2 and given by

$$\Delta L_{\tilde{W}} \le -\varpi_1 \left\|\tilde{W}_V\right\|^4 - \varpi_2 \left\|\tilde{W}_I\right\|^4 - \varpi_3 \left\|\tilde{W}_I\right\|^2 + \varepsilon_{\tilde{W}}. \tag{A.26}$$

Combining the individual first differences (A.23), (A.24), (A.25) and (A.26)

$$\Delta L_{cl} \le -\|\breve{x}\|^2 - \varpi_1 \left\|\tilde{W}_V\right\|^4 - \varpi_2 \left\|\tilde{W}_I\right\|^4 - \varpi_3 \left\|\tilde{W}_I\right\|^2 + B_{cl,2k}, \tag{A.27}$$

where $B_{cl,2k} = B_k^x + \varepsilon_{\tilde{W}}$ is a piece wise constant function.

From (A.27) it is evident that $\Delta L_{cl} < 0$ as long as

$$\left\| \breve{x} \right\| > B_{cl,2k} \quad \text{or}$$

$$\left\| \tilde{W}_V \right\| > \sqrt[4]{B_{cl,2k} / \varpi_1} \equiv B_{\tilde{W}_V k}^{JP} \quad \text{or}$$

$$\left\| \tilde{W}_I \right\| > \max\{\sqrt[4]{B_{cl,2k} / \varpi_2}, \sqrt{B_{cl,2k} / \varpi_3}\} \equiv B_{\tilde{W}_I k}^{JP}.$$

Hence, the system state, x, the identification error, e_I, last held system state, $\breve{x}$, NN weight estimation errors ,$\tilde{W}_V$, and $\tilde{W}_I$, are bounded.

Since, the system state x is ultimately bounded for $t_k > \bar{T}$ with an ultimate bound B_{ub}^x, as shown in Case I, $B_{cl,2k}$ in (A.27) converges to the ultimate value $B_{cl,2}^{ub} = B_{ub}^x + \varepsilon_{\tilde{W}}$ for all $t_k > \bar{T}$. Therefore, for all jump instants $t_k > \bar{T}$, the system state x, the last transmitted state $\breve{x}$, the NN weight estimation errors are, respectively, locally ultimately bounded as $\|x\| \le B_{ub}^x$, $\left\| \breve{x} \right\| \le B_{cl,2}^{ub}$, $\left\| \tilde{W}_V \right\| \le B_{\tilde{W}_V}^{JP,ub}$ and $\left\| \tilde{W}_I \right\| \le B_{\tilde{W}_I}^{JP,ub}$ where $B_{\tilde{W}_V}^{JP,ub} = \sqrt[4]{B_{cl,2}^{ub} / \varpi_1}$ and $B_{\tilde{W}_I}^{JP,ub} = \max\{\sqrt[4]{B_{cl,2}^{ub} / \varpi_2}, \sqrt{B_{cl,2}^{ub} / \varpi_3}\}$ for all $t_k > \bar{T}$. Therefore, the closed-loop nonlinear impulsive dynamical system state ζ locally UB for all $t_k > \bar{T}$ with an ultimate bound $\mu_c = \max\{B_{ub}^x, B_{cl,2}^{ub}, B_{\tilde{W}_V}^{JP,ub}, B_{\tilde{W}_I}^{JP,ub}\}$.

Consequently, from Both the cases the closed-loop nonlinear impulsive dynamical system state ξ locally UB with an ultimate bound $\mu = \max\{\mu_f, \mu_c\}$ for all $t_k > T$ or alternatively $t > \bar{T}$ for $T > \bar{T}$.

To show the convergence of the estimated value function near to the optimal value, consider the difference

$$\left\|V^{*}-\hat{V}\right\| \le \left\|\tilde{W}_V^T\phi(x)+\hat{W}_V^T\left(\phi(x)-\phi(\bar{x})\right)+\varepsilon_V\right\| \le \phi_M B_{\tilde{W}}^{ub}+\hat{W}_{V,M}C_{\phi}\sigma_{ETC,M}B_{ub}^{x}+\varepsilon_{V,M} \equiv B_V.$$

where $\sigma_{ETC,M}$, $\hat{W}_{V,M}=\max_{k\in\mathbb{N}}\left(\left\|\hat{W}_{V,k}\right\|\right)$ are the maximum value of the threshold over all flow interval, and B_V is a small positive constant. ■

Proof of theorem 2: Recall the closed-loop dynamics (37) of the event sampled system for $t_k < t \le t_{k+1}$. The upper bound of the system dynamics for k^{th} inter-sample time can be expressed as

$$\left\|\dot{x}\right\| \le K\left\|x\right\|+M_k, \;\; t_k < t \le t_{k+1}, \tag{A.28}$$

where $M_k = g_M\lambda_{\max}(R^{-1})\sigma_{I,M}\phi_M^{'}\left\|\hat{W}_{I,k}\right\|\left\|\hat{W}_{V,k}\right\|+\frac{1}{2}g_M\lambda_{\max}(R^{-1})\phi_M^{'}W_{V,M}(\sigma_{I,M}\left\|\tilde{W}_{I,k}\right\|+\bar{\varepsilon}_I)$

$+(1/2)g_M\lambda_{\max}(R^{-1})\sigma_{I,M}\phi_M^{'}\left\|\hat{W}_{I,k}\right\|\left\|\tilde{W}_{V,k}\right\|+(1/2)g_M^2\lambda_{\max}(R^{-1})\varepsilon_{V,M}^{'}$. Note that M_k is constant during the k^{th} flow period as the vale function NN weight estimation error $\tilde{W}_{V,k}$, $\hat{W}_{I,k}$ and value function NN weight estimates $\hat{W}_{V,k}$ remain constant during the the k^{th} flow period.

The derivative of the event sampling error e_{ET} can be expressed as

$$d\left(e_{ET}\right)/dt \le \left\|\dot{e}_{ET}\right\| = \left\|\dot{x}-\dot{\bar{x}}\right\| = \left\|\dot{x}\right\| \le -K\left\|x\right\|+M_k, \tag{A.29}$$

for $t_k < t \le t_{k+1}$, $\forall k = 1,2,\cdots$. By comparison lemma [18], the solution of the differential inequality (A.29) with initial condition $e_{ET}^{+}=0$ for $t=t_k$ is bounded above by

$$\left\|e_{ET}\right\| \le \int_{t_k^{+}}^{t}\exp\left(K\left(t-s\right)\right)M_k ds = M_k\left(\exp\left(K\left(t-t_k\right)\right)-1\right)/K, \tag{A.30}$$

for each $t_k < t \le t_{k+1}, \forall k = 1,2,\cdots$

To compute lower bound on the inter-sample times we consider the minimum value of the threshold over all flow intervals $t_k < t \le t_{k+1}$, $\forall k = 1, 2, \cdots$. The minimum threshold can be computed as

$$\sigma_{ETC,\min}\left(x, \hat{W}_I, \hat{W}_V\right) = \min_{k\in\mathbb{N}}\left(\sigma_{ETC}\left(x, \hat{W}_I, \hat{W}_V\right)\right) = \sigma_{ETC}\left(Q_{\min}, \hat{W}_{I,M}, \hat{W}_{V,M}\right) \tag{A.31}$$

where, $\hat{W}_{V,M} = \max_{k\in\mathbb{N}}\left(\left\|\hat{W}_{V,k}\right\|\right)$, $\hat{W}_{I,M} = \max_{k\in\mathbb{N}}\left(\left\|\hat{W}_{I,k}\right\|\right)$ are the maximum value of the NN weight estimates over all flow intervals $t_k < t \le t_{k+1}$ for all $k = 1, 2, \cdots$.

It is pertinent to mention here that $\hat{W}_{V,M}$ and $\hat{W}_{I,M}$ exist as the NN weight estimation errors are bounded as proves in Lemma 2. Further, the dead-zone operator (45) ensures a lower bound on the system state, i.e., B_{ub}^{x} for the sampling. Hence, by definition it holds that $Q(x) > Q_{\min} > 0$.

For a minimum inter-sample time, the event sampling condition (44), at next sampling instants satisfies

$$\left\|e_{ET}\left(t_{k+1}\right)\right\| = \sigma_{ETC,\min}. \tag{A.32}$$

Comparing (A.32) and (A.30) at t_{k+1} , it holds that

$$\sigma_{ETC,\min} \le \left(M_k / K\right)\left(\exp\left(K\left(t_{k+1} - t_k\right)\right) - 1\right). \tag{A.33}$$

Solving the inequality (A.33), one can reach at

$$t_{k+1} - t_k = \tau_k \ge \left(1/K\right)\ln\left(1 + \left(K/M_k\right)\sigma_{ETC,\min}\right) > 0, \ \forall k = 1, 2, \cdots \tag{A.34}$$

From (A.34), the inter-sample times $\tau_k > 0$ for $\forall k = 1, 2, \cdots$ as $\left(K/M_k\right)\sigma_{ETC,\min} > 0$. Consequently, the minimum inter-sample time

$$\tau_{\min} \ge \min_{k\in\mathbb{N}}\left\{\left(1/K\right)\ln\left(1 + \left(K/M_k\right)\sigma_{ETC,\min}\right)\right\} > 0. \qquad \blacksquare$$

SECTION

2. CONLUSIONS AND FUTURE WORK

In this dissertation, event-sampled deterministic and stochastic Q-learning and adaptive dynamic programming techniques were developed for linear and a class of nonlinear systems for both in discrete and continuous time domain. The application of the designs for the NCS with time-varying delays and packet losses are also included. The system dynamics both in the linear and the nonlinear cases were considered completely uncertain. The universal approximation property of the neural network (NN) was revisited and event sampled approximation was derived. The event sampled approximation was used to estimate the deterministic and stochastic Q-function (for NCS) for linear systems and approximate the system dynamics, value function and optimal control input for nonlinear systems. The aperiodic transmission and controller execution instants were determined by designing novel adaptive event sampling conditions. The event sampling conditions orchestrated the sampling and transmission instants to achieve the accuracy in estimation/approximation and control performance with effective resource utilization.

2.1 CONCLUSIONS

In the first paper, the event sampled optimal adaptive regulation of a linear discrete time system using both state and output feedback is solved in a forward in time manner without requiring the knowledge of the system dynamics. Event driven Q-learning techniques were developed both for state and output feedback to design the optimal control policies. The designed event sampled optimal adaptive control policies were able to regulate the system states in both the cases with a reduced number of controller

executions. The adaptive event sampling conditions found to generate required number of event sampled instants to achieve desired estimation accuracy and, hence, optimality. It was observed that the output feedback design resulted in more number of controller execution when compared to the state feedback. This was due to the additional uncertainty introduced by the adaptive observer used to reconstruct the system state. In addition, the event sampling condition turned out to be a function of estimated Q-function parameter ensuring the desired performance and stability. It was further observed that the initial Q-function parameters and the learning gain of the parameter tuning law affect the number of event sampled instants. The application of this technique for NCS with network induced time varying delays and random packet losses found to regulate the system for delays longer than a sampling interval. The redesigned adaptive event sampling condition for the stochastic system and tuning law ensured the asymptotic convergence in the mean of the closed-loop system with 56% average saving in communication.

On the other hand, for the case of nonlinear systems, in Paper II, a NN based adaptive state estimator was employed as a model. The event sampled NN based approximation and weight update scheme approximated the unknown nonlinear functions with a small bounded error. The event sampled instants were occurred frequently during the initial learning phase, but the inter-sample times were increased with the convergence of the NN weight estimates to their respective target values. Further, it was observed that the change in the NN weight initialization and learning gains for the weight update schemes affect the number of controller update. These results were validated with the numerical examples. The dead zone operator, used to stop the unnecessary triggering

once the system state is in the ultimate bound, found to further reduce the redundant computation. The introduction of a model increased the computation when compared to model free schemes whereas it is found to be more effective in reduction of event sampled or transmission instants and best suitable for NCS.

An event sampled near optimal adaptive regulator was proposed in Paper III. The actor-critic frame work, to solve the finite horizon optimal control problem in a forward in time manner, was redesigned with event-sampled feedback information; leading to an event-driven adaptive dynamic programming. The near optimality is achieved in a finite time with complete unknown system dynamics. The novel identifier structure proposed to approximate the system dynamics with intermittent update at the event sampled instants performed satisfactorily. The aperiodic update scheme at the event sampled instants determined by the adaptive event sampling condition drove the NN weight estimation errors within a small bound. With an explicit formula, the existence of the non-triviality of the inter-sampled times were proven and corroborated by the simulation results. An event-sampled stochastic ADP scheme was also developed to overcome the time-varying network induced delays and packet losses as an application for the proposed design. The stochastic optimal controller performed satisfactorily for delays more than a sampling time.

The fourth paper presented a continuous time event-based control using approximate feedback linearization. The novel NN weight update law as a jump in the weights at the event sampled instants able to approximate the control input with aperiodic update. It is observed that the initial NN learning phase plays a key role in ensuring minimum inter-event time. In the final paper the event sampled continuous time ADP

scheme guaranteed near optimality with a reduction in computation. The approximation of the value function with event-sampled information was found satisfactory with a near optimal performance. Similar to the previous paper, a larger number of triggers were observed in the initial learning phase and therefore a proper initialization of the weights are necessary to ensure the lower bound on the inter-sample times.

2.2 FUTURE WORK

As part of the future work, the optimization of the event sampled instants could be considered. This needs a redefinition of the performance index by penalizing the transmission instants which is not considered in this thesis. Although a few results are available in the literature utilizing certainty equivalence principle, this problem in an adaptive dynamic programming frame work will be a challenging one for a forward in time solution. However, this will increase the effectiveness of resource utilization for the networked control systems.

On the other hand, the event sampled control is best suitable for spatially distributed systems. The nationwide pervasive distributed systems, such as, electrical power grid, transportation system, formation control of mobile robots to name a few, need large amount of computational power and communication network bandwidth. Although, there are quite a few results available in the literature, event sampled optimal control of uncertain distributed interconnected system is a perspective area and can be part of future research. Application of the event based optimal control to formation control, consensus based control can be another future area of research.

Self-triggered control which is the counter part of the event-triggered control could be an area of future research to explore the optimal event design schemes without continuous measurement. This would be a more interesting problem for systems with uncertain dynamics.

APPENDIX A

REGULATION OF LINEAR NETWROKED CONTROL SYSTEMS BY USING EVENT SAMPLED Q-LEARNINGAND DYNAMIC PROGRAMMING

A. REGULATION OF LINEAR NETWORKED CONTROL SYSTEMS BY USING EVENT SAMPLED Q-LEARNING AND DYNAMIC PROGRAMMING

Avimanyu Sahoo and S. Jagannathan

Abstract *— In this paper, the optimal regulation of networked control systems (NCS) in the presence of time-varying delays and random packet losses is presented by using event sampled state and input vector. A stochastic optimal regulator is designed using adaptive dynamic programming and Q-learning technique with event sampled feedback information. The Q-function parameters are tuned at the event sampled instants in a aperiodic manner with a novel parameter tuning law. An adaptive event sampling condition is derived analytically to determine the event sampled instants. This adaptive sampling condition not only maintains stability and saves communication but also facilitates parameter estimation with event sampled state and input information. The asymptotic stability in the mean of the closed-loop system is demonstrated through Lyapunov technique. A condition for non-trivial inter-sample times is derived. Finally, simulation results are included to substantiate the analytical design.*

Keywords - Q-learning, event sampled control, adaptive dynamic programming, adaptive control, optimal control.

1. INTRODUCTION

Networked control systems (NCS) [1]-[4] are gaining popularity in recent years due to their reduced complexity and distributed architecture. However, the communication network introduces various imperfections such as time-varying delays, random packet losses, and quantization errors. These network artifacts deteriorate NCS performance and may jeopardize the stability of the closed-loop system [2] with a traditional controller. In the recent past, an ample amount of research has been carried out for studying the stability of NCS [1]-[4] in the presence of the network imperfections. From the optimal control point of view for NCS, stochastic Riccati equation based approach [2][4] is used and solved backward in time.

In contrast, forward-in-time solution of the optimal policy is obtained by using reinforcement learning [5]-[7], adaptive dynamic programming (ADP) [8], and Q-learning [2]-[3] techniques. These schemes use policy or value iteration to solve the Bellman [9] or Hamilton-Jacobi-Bellman (HJB) equation [8] required for computing optimal control policy. Instead of using these computational intensive iterative techniques, the authors in [3]-[4] presented a time-based adaptive Q-learning approach for NCS by using the time history of the temporal difference (TD) or Bellman error to estimate the Q-function. In all the schemes [1]-[4], a fixed sampling interval is utilized to transmit the feedback information from the system to the controller thus requiring a large network bandwidth.

To alleviate the problem of higher bandwidth usage, an alternate framework referred to as event sampled or triggered control [11]-[13] is introduced in the recent times. This control paradigm reduces both the network traffic and computation by

orchestrating the transmissions and controller executions using a state dependent sampling scheme. The main idea behind the event-triggered control design is the selection of aperiodic sampling/transmission instants or simply referred as *event sampled instants* without sacrificing stability and performance. The event sampled optimal control of NCS is studied by the authors in [13]-[14]. The optimal control design still uses the backward-in-time solution of the Riccati equation (RE) with known system dynamics.

In this paper, we present an optimal adaptive regulation of NCS represented as an uncertain linear continuous-time system in the presence of the time-varying delays and packet losses with event sampled state and input vector. The system state vector is sampled periodically by the sensor whereas the feedback signals are transmitted only at event sampled instants. The optimal regulator is designed by using novel stochastic Q-learning and ADP [8] with event sampled state and input vector without using the knowledge of system dynamics.

An adaptive estimator is designed to estimate the action dependent value or Q-function parameters which are tuned at the aperiodic event sampled instants. In contrast to traditional event-triggered control [11]-[12], an adaptive event sampling condition is designed to determine the event sampled instants which facilitates the estimation of the Q-function parameters while retaining the advantages of the traditional event-triggered control [11]-[13] in terms of resource saving, stability and performance. A condition to show the existence of the non-trivial inter-sample times is presented. However, the optimization of the event sampled instants is not attempted.

The remaining of the paper is organized as follows. Section 2 gives a brief background of NCS and formulates the problem. The design procedure and simulation results are presented in Section 3 and Section 4, respectively. Section 5 presents the conclusion. The Appendix gives the proof for the theorems and the corollaries.

2. BACKGROUND AND PROBLEM FORMULATION

In this section, a brief background on stochastic Q-learning is discussed. Then, the problem of event sampled Q-learning is addressed by investigating the effect of the aperiodic transmissions on estimation and stability.

2.1 SYSTEM FORMULATION AND PERIODIC Q-LEARNING

Consider the NCS shown in Figure 1 and represented by a linear time invariant (LTI) continuous-time system given by

$$\dot{x}(t) = Ax(t) + Bu(t), \tag{1}$$

where $x(t) \in \Re^n$ and $u(t) \in \Re^m$ are the system state and the control input vectors, respectively, with $A \in \Re^{n\times n}$ and $B \in \Re^{n\times m}$ being unknown system matrices. Before proceeding further the main assumption on the NCS are introduced.

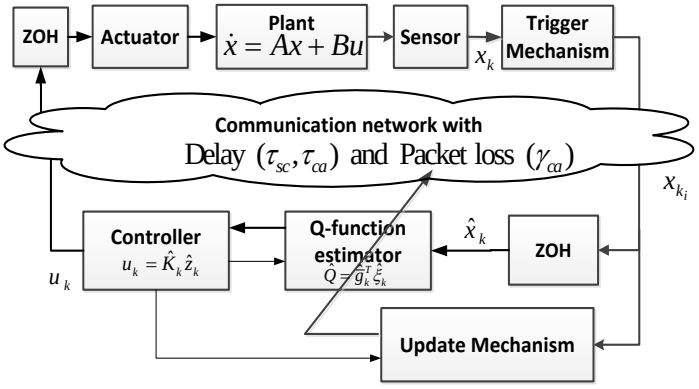

Figure 1. Structure of the event sampled networked control system.

Assumption 1: (a) The system (1) is controllable and the order of system is known [3]. The state vector is measurable and the control coefficient matrix B satisfies $\|B\| \le B_{\max}$

where $B_{\max} > 0$ being a known constant. (b) The sensor is time-driven and it samples the state vector with a fixed periodic sampling time T_s [1]-[4]. (c) The networked induced time-varying delays from the sensor to the controller, $\tau_{sc}(t)$, and controller to the actuator $\tau_{ca}(t)$, respectively, satisfy $\tau_{sc}(t) < \Delta_{sc}$ and $\tau_{ca}(t) < \bar{d}T_s$ where $\Delta_{sc} \leq T_s$ is a constant skew between the sensor and controller sampling instants [1] and $\bar{d}$ is a positive integer. (d) The packet losses from the sensor to the controller, γ_{sc} is negligible [3], while from the controller to the actuator γ_{ca} follows the Bernoulli distribution.

The system (1) uses a discrete time controller due to the packet switched network in the feedback loop. The discretized version of the system (1) after incorporating the delays and the packet losses, similar to that in [1]-[4], can be represented as

$$z_{k+1} = A_{z,k} z_k + B_{z,k} u_k, \tag{2}$$

where $A_{z,k} \in \Re^{(n+\bar{d}m)\times(n+\bar{d}m)}$ and $B_{z,k} \in \Re^{(n+\bar{d}m)\times m}$ are the transformed system matrices [3] given by

$$A_{z,k} = \begin{bmatrix} A_s & \gamma_{ca,k-1}B_{1,k} & \cdots & \gamma_{ca,k-l}B_{l,k} & \cdots & \gamma_{ca,k-\bar{d}}B_{\bar{d},k} \\ 0 & 0 & \cdots & \cdots & \cdots & 0 \\ 0 & I_m & 0 & \cdots & \cdots & 0 \\ \vdots & 0 & I_m & 0 & \cdots & 0 \\ \vdots & \vdots & 0 & \ddots & \cdots & \vdots \\ 0 & 0 & \cdots & \cdots & I_m & 0 \end{bmatrix} \text{ and } B_{z,k} = \begin{bmatrix} \gamma_{ca,k}B_{0,k} \\ I_m \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}$$

The vector $z_k = [x_k^T \quad u_{k-1}^T \quad u_{k-2}^T \quad \cdots \quad u_{k-\bar{d}}^T]^T \in \Re^{n+\bar{d}m}$ is the augmented state vector with $x_k \in \Re^n$ and $u_k \in \Re^m$ are the discrete-time state and the control input vectors at time kT_s. The matrices $A_s = e^{AT_s}$ and $B_{l,k} = \int_{t_l^k}^{t_{l-1}^k} e^{A(T_s - s)} B ds$ represent the discretized system matrices

and t_l^k, $l = 0,1,2,\cdots,\bar{d}$ is the time instant after kT_s when the delayed control inputs are applied to the plant. The variable $\gamma_{ca,k}$ is the packet loss indicator given by

$$\gamma_{ca,k} = \begin{cases} I^{n\times n}, & k^{th}\text{control input is recieved} \\ 0^{n\times n}, & k^{th}\text{ control input is not received}, \end{cases} \tag{3}$$

where $I^{n\times n}$ is the identity matrix and $0^{n\times n}$ is a matrix with all elements zero. The augmented system matrices $A_{z,k}$ and $B_{z,k}$ are stochastic and time-varying. The integer $\bar{d}$ is selected such that the pair $(A_{z,k}, B_{z,k})$ [4] is controllable. From Assumption 1(a) it is clear that $\|B_{z,k}\| \le B_{z,\max}$ where $B_{z,\max} > 0$ a known constant.

From the optimal control point of view, consider the stochastic cost function [3] for the system (2) given by

$$V_k = \underset{\tau,\gamma}{E}\left\{\sum_{j=k}^{\infty} r(z_j, u_j)\right\}, \tag{4}$$

where $r(z_k, u_k) = z_k^T P_z z_k + u_k^T R_z u_k$ is the cost-to-go at the time instant k. The matrices P_z and R_z, respectively, are positive semi-definite and positive definite penalty matrices of appropriate dimensions and $\underset{\tau,\gamma}{E}\{\bullet\}$ is the expectation operator (mean value) with $\tau = \tau_{sc}$ and $\gamma = \gamma_{ca}$ for brevity. The optimal control input can be computed online and in a forward-in-time manner without knowledge of $A_{z,k}$ and $B_{z,k}$ by using the stochastic Q-learning based scheme discussed in [2]-[4]. A brief back ground is discussed here.

The stochastic optimal action dependent value function or the Q-function [3] is given by

$$Q^*(z_k, u_k) = \underset{\tau,\gamma}{E}\{(r(z_k, u_k) + V_{k+1})\} = [z_k^T \quad u_k^T]\underset{\tau,\gamma}{E}\{G_k\}[z_k^T \quad u_k^T]^T = w_k^T \bar{G}_k w_k, \tag{5}$$

where V_{k+1} is the cost for time $k+1$ onwards, $w_k = [z_k^T \quad u_k^T]^T \in \Re^{l_{mn}}$, $l_{mn} = n + (\bar{d}+1)m$ and $\bar{G}_k \in \Re^{q\times q}$, $q = n + (\bar{d}+1)m$, is the Q-function parameter matrix given by

$$\bar{G}_k = \underset{\tau,\gamma}{E}\{G_k\} = \begin{bmatrix} P_z + \underset{\tau,\gamma}{E}\{A_{z,k}^T S_{k+1} A_{z,k}\} & \underset{\tau,\gamma}{E}\{A_{z,k}^T S_{k+1} B_{z,k}\} \\ \underset{\tau,\gamma}{E}\{B_{z,k}^T S_{k+1} A_{z,k}\} & R_z + \underset{\tau,\gamma}{E}\{B_{z,k}^T S_{k+1} B_{z,k}\} \end{bmatrix} \equiv \begin{bmatrix} \bar{G}_k^{xx} & \bar{G}_k^{xu} \\ \bar{G}_k^{ux} & \bar{G}_k^{uu} \end{bmatrix} \tag{6}$$

The optimal control input from (6) is given by

$$u_{p,k}^* = -(\bar{G}_k^{uu})^{-1}\bar{G}_k^{ux} z_k = -K_k^* z_k, \tag{7}$$

where $K_k^* = (\bar{G}_k^{uu})^{-1}\bar{G}_k^{ux} = [R_z + \underset{\tau,\gamma}{E}(B_{z,k}^T S_{k+1} B_{z,k})]^{-1} \underset{\tau,\gamma}{E}(B_{z,k}^T S_{k+1} A_{z,k})$.

With the following standard assumption, the Q-function parameters are estimated online to execute the control (7).

Assumption 2 [3]: The matrix $\bar{G}_k$ is slowly time-varying and can be expressed as a linear in the unknown parameters.

Then, the parametric form for (5) can be represented as

$$Q^*(z_k, u_k) = \bar{g}_k^T \xi_k, \tag{8}$$

where $\bar{g}_k \in \Re^{l_g}$, $l_g = q(q+1)/2$, is the vector form of $\bar{G}_k$ as in [3], and $\xi_k = w_k \otimes w_k$ is the regression vector with $\otimes$ denotes the Kronecker product. The estimated value of the stochastic Q-function (8) can be expressed as

$$\hat{Q}(z_k, u_k) = w_k^T \hat{\bar{G}}_k w_k = \hat{\bar{g}}_k^T \xi_k, \tag{9}$$

where $\hat{\bar{G}}_k = \begin{bmatrix} \hat{\bar{G}}_k^{xx} & \hat{\bar{G}}_k^{xu} \\ \hat{\bar{G}}_k^{ux} & \hat{\bar{G}}_k^{uu} \end{bmatrix}$ is the estimates of $\bar{G}$ and $\hat{\bar{g}}_k \in \Re^{l_g}$ is the vector form of $\hat{\bar{G}}_k$. It is known that the optimal cost function $V_k^* = Q^*(z_k, u_{p,k}^*)$ when the control input is optimal.

By Bellman principle of optimality, using the parametric form (8), the value function satisfies

$$0 = \underset{\tau,\gamma}{E}\{r(z_k, u_k)\} + V_{k+1} - V_k = \underset{\tau,\gamma}{E}\{r(z_k, u_k)\} + \overline{g}_k^T \Delta\xi_k, \tag{10}$$

where $\Delta\xi_k = \xi_{k+1} - \xi_k$. The Bellman or TD error with the estimated Q-function (9) can be expressed as

$$e_{V,k} = \underset{\tau,\gamma}{E}\{r(z_k, u_k)\} + \hat{\overline{g}}_k^T \Delta\xi_k . \tag{11}$$

It was shown in [3]-[4] that by using an initial admissible policy, the optimal control policy can be attained by adjusting the Q-function parameters and using the augmented time history of the Bellman error (11). This relaxes the value and policy iterations. Here periodically sampled system state vector, x_k, is transmitted to the controller for implementing the scheme which consumes significant network bandwidth. The problem of event sampled intermittent transmission is discussed next.

2.2 PROBLEM FORMULATION

In this paper, our main objective is to transmit the system state vector x_k and execute and transmit the control policy u_k at the event sampled instants. In addition, the Q-function parameters are estimated with the constrained information to achieve optimality. Let a subsequence $\{k_i\}_{i=1}^{\infty}$ of time instants $k \in \mathbb{N}$ be the event sampled instants with initial transmission at $k_0 = 0$. Then, the transmitted state vector held at the controller till the next transmission denoted by $\hat{x}_k$ is given by

$$\hat{x}_k = x_{k_i}, \quad k_i \le k < k_{i+1}, \forall i = 1, 2, \cdots. \tag{12}$$

Thus, the event sampled augmented state vector $\hat{z}_k$ at the controller is defined as

$$\hat{z}_k = z_{k_i},\ k_i \le k < k_{i+1},\ \forall i = 1,2,\cdots, \tag{13}$$

where $z_{k_i} = [x_{k_i}^T \quad u_{k_{i-1}}^T \quad u_{k_{i-2}}^T \quad \cdots \quad u_{k_{i-\bar{d}}}^T]^T$, $\forall i = 1,2,\cdots$ is the augmented state formed by storing the previous values. The error introduced by the event sampled transmission referred to as event sampling error is given by

$$e_{ET,k} = z_k - \hat{z}_k,\ k_i \le k < k_{i+1},\ \forall i = 1,2,\cdots. \tag{14}$$

The event sampled instants k_i, $\forall i = 1,2,\cdots$ are determined by a state dependent event sampling condition evaluated at each sensor sampling instant k at the trigger mechanism. A transmission decision is made only at the violation of the condition. This enables the control update and resets the error, $e_{ET,k}$, to zero for the next cycle of operation.

Now, the optimal control input (7) with event sampled state vector (13) becomes

$$u_{e,k}^* = -K_k^* \hat{z}_k = -K_k^* z_k + K_k^* e_{ET,k},\ k_i \le k < k_{i+1}. \tag{15}$$

The estimation of the Q-function in a parametric form (9) with event sampled state vector (13) is rewritten as

$$\hat{Q}(\hat{z}_k, u_k) = \hat{w}_k^T \hat{\bar{G}}_k \hat{w}_k = \hat{\bar{g}}_k^T \hat{\xi}_k,\ k_i \le k < k_{i+1},\ \forall i = 1,2,\cdots, \tag{16}$$

where $\hat{\xi}_k = \hat{w}_k \otimes \hat{w}_k$ is the event sampled regression vector with $\hat{w}_k = [\hat{z}_k^T \quad u_k^T]^{\mathrm{T}}$. The event sampled Bellman error (11) by using (12) becomes

$$e_{V,k} = \underset{\tau,\gamma}{E}\{r(\hat{z}_k, u_k)\} + \hat{\bar{g}}_k^T \Delta \hat{\xi}_k,\ k_i \le k < k_{i+1}, \forall i = 1,2,\cdots. \tag{17}$$

The additional error introduced due to event sampled transmission can be seen by rewriting (17) in the form of (11) to get $e_{V,k} = \underset{\tau,\gamma}{E}\{r(z_k, u_k)\} + \hat{\bar{g}}_k^T \Delta \xi_k + \Xi(z_k, e_{ET,k})$. The

error $\Xi(z_k, e_{ET,k}) = \underset{\tau,\gamma}{E}\{r(\hat{z}_k, u_k) - r(\hat{z}_k, u_k, e_{ET,k})\} + \hat{\bar{g}}_k^T(\Delta\hat{\xi}_k - \Delta\xi_k)$ is a function of the event sampling error both in the cost-to-go and the regression vector. Further, since the Bellman error (17) gets updated only at the event sampled instants, the Q-function parameters must only be updated at that time. This reduces the frequency of the parameter update when compared to the traditional adaptive control [15]. Since the Bellman error is driven by event sampling error, the proposed technique is referred to as event driven dynamic programming.

From the above discussion, for an optimal policy the event sampling condition needs to be designed in such a way that the estimated Q-function parameters converge to the optimal values while keeping the transmission small. Since the event sampling error is driving the estimation, a smaller error will increase the accuracy of estimation, transmission and computation. Therefore, a properly designed event sampling condition is necessary to create suitable number of transmissions in order to minimize this error as presented in the next section.

3. EVENT SAMPLED OPTIMAL REGULATOR DESIGN

The structure shown in Figure 1 is used for the proposed design. In contrast with the traditional event-triggering conditions [11]-[12], the threshold in the proposed event sampling condition is adjusted with the update in the estimated Q-function parameters. This leads to a transmission scheme based on both the parameter estimation error and system state. On the other hand, once the parameters converged close to target values, it becomes the traditional event-triggering condition [11]-[12].

The challenge in implementing the adaptive event sampling condition is the transmission of the estimated Q-function parameters between the trigger mechanism and controller which will require a large bandwidth. This can be avoided by using a mirror Q-function estimator at the trigger mechanism to estimate the Q-function locally provided the information at both sides of the network (trigger mechanism and controller) are same. Thus, the mirror Q-function estimator is designed to operate in synchronism with the one at the controller. In other words, the mirror and the actual Q-function estimators are initialized with the same initial conditions and get updated at the event sampled instants. As we will see later, the update of the Q-function parameter estimates require the state vectors, x_{k_i} and x_{k_i-1}, $\forall i = 1, 2, \cdots$, it is proposed that they are packetized together and transmitted to the controller at the event sampled instants.

Remark 1: From Assumption 1, the delays between the sensors to controller satisfy $\tau_{sc} \leq \Delta_{sc}$ and the packet losses are considered negligible. Thus, the mirror and the actual Q-function estimator use the same state information for updates. The addition of a mirror increases the computation when compared to a single estimator at the controller but the overall computation is reduced due to event sampled implementation.

3.1 EVENT SAMPLED Q-LEARNING AND CONTROLLER DESIGN

Consider the event sampled estimation of the Q-function in (16) and the Bellman error in (17). The Q-function estimated parameters will be updated only at the event sampled instants with $\hat{z}_k = z_k$ for $k = k_i$, $\forall i = 1, 2, \cdots$ using the Bellman error. The Bellman error (17) at the event sampled instants becomes

$$e_{V,k} = \underset{\tau,\gamma}{E}\{r(z_k, u_k)\} + \hat{\bar{g}}_k^T \Delta\xi_k, \ k = k_i, \forall i = 1, 2, \cdots. \tag{18}$$

An augmented Bellman error, similar to [3], using the time history can be written as

$$E_{V,k} = \Pi_k + \hat{\bar{g}}_k^T Z_k, \ k = k_i, \forall i = 1, 2, \cdots, \tag{19}$$

where $\Pi_k = [\underset{\tau,\gamma}{E}\{r(z_{k_i}, u_{k_i})\} \ \underset{\tau,\gamma}{E}\{r(z_{k_{i-1}}, u_{k_{i-1}})\} \ \cdots \ \underset{\tau,\gamma}{E}\{r(z_{k_i-1-v}, u_{k_{i-1-v}})\}] \in \Re^{1\times v}$ and $Z_k = [\Delta\xi_{k_i} \ \ \Delta\xi_{k_{i-1}} \ \cdots \ \Delta\xi_{k_{i-1-v}}] \in \Re^{l_g \times v}$ is the history of the cost-to-go and regression vector, respectively. The difference between the augmented error (19) and that in [3] is that (19) uses the histories only at the event sampled instants. It is clear that convergence of $E_{V,k}$ to zero ensures the convergence of $e_{V,k}$. The length of the time history v is determined based on the designer's experience and $v \le l_g - 1$ found to be suitable for the estimation during the simulation study. A larger size of the history information will lead to a faster convergence.

The tuning law to update the Q-function parameter estimates can be selected as

$$\hat{\bar{g}}_k = \begin{cases} \hat{\bar{g}}_{k-1} - \dfrac{\alpha_V Z_{k-1} E_{V,k-1}^T}{\left\| I + Z_{k-1}^T Z_{k-1} \right\|}, & k = k_i, \\ \hat{\bar{g}}_{k-1}, & k_{i-1} < k < k_i, \end{cases} \tag{20}$$

where $\alpha_v > 0$ is the adaptive gain parameter. The tuning law (20) is aperiodic due to the event sampled instants. This reduces the computation when compared to a periodically sampled adaptive control [15].

Remark 2: The computation of the regression matrix Z_{k_i-1} in (20) at the event sampled instants requires the augmented state z_{k_i} and z_{k_i-1} which is formed at the controller using the received state vectors, x_{k_i} and x_{k_i-1}. Therefore, it is proposed that both the state vectors are packetized together and transmitted at the event sampled instants.

By using the estimated Q-function parameters, the event sampled estimated control policy can be expressed as

$$u_k = -\hat{K}_k \hat{z}_k,\ k_i \le k < k_{i+1},\ \forall i = 1, 2, \cdots, \tag{21}$$

where $\hat{K}_k = (\hat{\bar{G}}_k^{uu})^{-1} \hat{\bar{G}}_k^{ux}$ with $\hat{\bar{G}}_k^{ux}$ and $\hat{\bar{G}}_k^{uu}$ defined in (9). Denoting the Q-function parameter estimation error $\tilde{\bar{g}}_k = \bar{g}_k - \hat{\bar{g}}_k$, the dynamics, from (20), can be represented as

$$\tilde{\bar{g}}_{k+1} = \begin{cases} \tilde{\bar{g}}_k + \dfrac{\alpha_V Z_k E_k^{V^T}}{\left\| I + Z_k^T Z_k \right\|}, & k = k_i, \\ \tilde{\bar{g}}_k, & k_i < k < k_{i+1}. \end{cases} \tag{22}$$

Remark 3: It is necessary that the regression matrix Z_k must satisfy the persistency of the excitation (PE) condition [15] for the convergence of the estimation error $\tilde{\bar{g}}_k$ to zero. This further implies the regression vector ξ_k must satisfy the PE condition. The PE condition can be enforced on the regression vector ξ_k by adding an exploration noise to the control input as discussed in [3].

The event sampled Bellman error can alternatively be represented in terms of the parameter estimation error $\tilde{\bar{g}}_k$ by subtracting (10) from (18) and found to be $e_{V,k} = -\tilde{\bar{g}}_k^T \Delta\xi_k,\ k = k_i$. Thus, the augmented Bellman error can be written as

$$E_{V,k} = -\tilde{\bar{g}}_k^T Z_k,\ k = k_i. \tag{23}$$

We will use this alternate expression while carrying out the proof for the following theorem.

Theorem 1: Consider the NCS dynamics (1), represented as an augmented system (2) along with the Q-function estimator (16) and the parameter tuning law (20). Let the Assumptions 1 and 2 hold and the initial parameters $\hat{\bar{g}}_0 \in \Omega_g$ with Ω_g a compact set. Suppose u_0 be the initial admissible control policy and the regression vector ξ_k satisfies the PE condition. Then, for an adaptive gain $0 < \alpha_V < 2$, the Q-function parameter estimation error $\tilde{\bar{g}}_k$ converges asymptotically to zero in the mean when the trigger instants $k_i \to \infty$, or , alternatively, the time $k \to \infty$.

Proof: Refer to the Appendix.

3.2 EVENT SAMPLING CONDITION AN D CONVERGENCE ANALYSIS

The closed-loop dynamics of the event sampled system by using (2) and (22) is given by

$$z_{k+1} = A_{z,k} z_k - B_{z,k} \hat{K}_k z_k + B_{z,k} \hat{K}_k e_{ET,k},\ k_i \le k < k_{i+1}.$$

Defining the control gain error $\tilde{K}_k = K_k^* - \hat{K}_k$ and using (7) the closed-loop dynamics can be rewritten as

$$z_{k+1} = A_{z,k} z_k + B_{z,k} u_{p,k}^* + B_{z,k} \tilde{K}_k z_k + B_{z,k} \hat{K}_k e_{ET,k},\ k_i \le k < k_{i+1}. \tag{24}$$

Next we introduce the following adaptive event sampling condition. Consider the event sampling error (14). The system states are transmitted when the condition

$$\left\| e_{ET,k} \right\| \le \sigma_{ET,k} \left\| z_k \right\|, \tag{25}$$

is violated where $\sigma_{ET,k} = \sqrt{\Gamma\left(1-3\mu\right)\Big/ 3B_{z,\max}^2 \left\| \hat{K}_k \right\|^2}$ is the adaptive threshold coefficient , $0 < \Gamma < 1$ and $\hat{K}_k$ is defined in (22). The constant μ satisfies $\left\| \underset{\tau,\gamma}{E}\{A_{z,k} z_k + B_{z,k} u_{p,k}^*\} \right\|^2 \le \mu \left\| z_k \right\|^2$ as discussed in [3]. Further, it was found from Theorem 2 that $0 < \mu < 1/3$ will ensure asymptotic stability in the mean of the event sampled system. To ensure the estimated value of $\|\hat{K}_k\|$ is not zero while evaluating the trigger condition (25), the previous nonzero value is used when the estimated value becomes zero.

The asymptotic stability in the mean [17] for the closed loop system (24) is shown by considering a single Lyapunov function candidate, L_k, for both event sampled instants and inter-sample times. As the parameters and the control policy are only updated at the event sampled instants and held during the inter-sample times, the Lyapunov function may not decrease monotonically during both the cases. This is not necessary as discussed in [12]. We only need to show the existence of a piecewise continuous function $h(k) \in \Re^+$, such that

$$h_k \ge L_k \text{ and } \lim_{k\to\infty} h_k = 0,\ k \in \mathbb{N}, \tag{26}$$

holds as illustrated in Figure 3 given below.

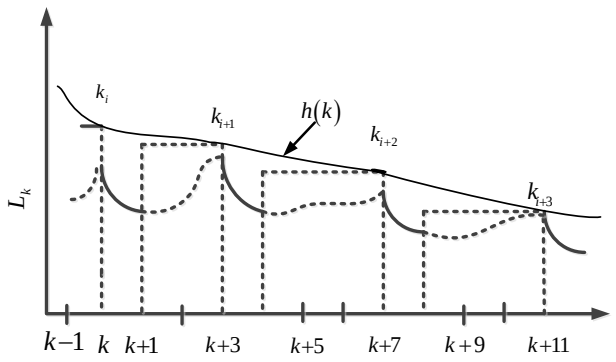

Figure 2. Evolution of the Lyapunov function during the event sampled instants and inter-sample times and the upper bound function.

Theorem 2: Consider the uncertain LTI discrete-time system (2) with the Q-function estimator (16) and the state feedback controller (21). Suppose the Assumptions 1 and 2 hold and the regression vector, ξ_k, satisfies the PE condition. Let the initial parameter estimate $\hat{\bar{g}}_0 \in \Omega_g$. Given an initial admissible control policy $u_0 \in \Omega_u \subset \Re^m$ and adaptive gain parameter satisfying $0 < \alpha_V < 2$, the closed-loop event sampled system (24), with the event sampling condition (29) and update law (20), is asymptotically stable in the mean as $k_i \to \infty$. In addition, the control input $u_k \to u_{e,k}^*$ as $k_i \to \infty$ or alternatively, $k \to \infty$.

Proof: Refer to the Appendix.

Corollary 1: Consider the NCS (2) with the Q-function estimator (16) and the state feedback controller (21). Then closed-loop event sampled system (24) is asymptotically stable in the mean with the event sampling condition given by

$$\left\| e_{ET,k} \right\| \le \left(\sigma_{ET,k} / (1 + \sigma_{ET,k}) \right) \left\| \hat{z}_k \right\|. \tag{27}$$

Proof: Refer to the Appendix.

The event sampling condition is evaluated at every time instants k with a fixed sensor sampling time T_s. Therefore, the minimum time between two consecutive event sampled instants is T_s. Alternatively, $\delta k_{\min} = \min_{i \in \mathbb{N}} \{k_{i+1} - k_i\} = 1$. A condition to achieve the non-trivial inter-sample times i.e., $\delta k_i = k_{i+1} - k_i > 1$, to save the communication and computational load, is presented next.

Theorem 3: Given the hypothesis of Theorem 2 with the event sampling condition (27), the inter-sample times $\delta k_i = k_{i+1} - k_i$ implicitly defined by (27) satisfies

$$\delta k_i \geq \ln\left(1 + (1/M_i)(F-1)\sigma_i\right) \Big/ \ln\left(F\right), \ \forall i = 1, 2, \cdots, \tag{28}$$

where $\sigma_i = \sigma_{ET,k_i} / (1 + \sigma_{ET,k_i})$ for the i^{th} inter-sample time. Further, $F = \sqrt{\mu} + B_{z,\max} K^*_{\max}$ and $M_i = (\sqrt{\mu} + B_{z,\max} \left\| \tilde{K}_{k_i} \right\| + 1)$ with $K^*_{\max}$ is the maximum value of the optimal control gain matrix. Further the inter-sample times δk_i become non-trivial when $\sigma_i / M_i > 1$.

Proof: Refer to the Appendix.

Remark 4: The function M_i and the threshold coefficient σ_{ET,k_i} depend on the control gain estimation error $\tilde{K}_k$ via the relation $\tilde{K}_k = K^*_k - \hat{K}_k$. Hence, the inter-sample times δk_i in (28) are a function of $\tilde{K}_k$ or $\tilde{\bar{g}}_k$. It is clear that the convergence of $\tilde{\bar{g}}_k$ close to zero, as proven Theorem 1, will satisfy the non-triviality condition. This further implies that the number of triggers will depend on the initial Q-function parameters and the adaptive learning gains.

4. SIMULATION RESULTS

The bench-mark example of batch reactor [2] is presented as an example in this section whose continuous-time dynamics are represented by

$$\dot{x} = \begin{bmatrix} 1.38 & -0.2077 & 6.715 & -5.676 \\ -0.5814 & -4.29 & 0 & 0.675 \\ 1.067 & 4.273 & -6.654 & 5.893 \\ 0.048 & 4.273 & 1.343 & -2.104 \end{bmatrix} x + \begin{bmatrix} 0 & 0 \\ 5.679 & 0 \\ 1.136 & -3.146 \\ 1.136 & 0 \end{bmatrix} u.$$

The following parameters were chosen for simulation. The initial vales were as chosen as $z_0 = \begin{bmatrix} 0.2 & 1 & -3 & 0.5 \end{bmatrix}^T$ and $\hat{\bar{g}}_0$ is random from a uniform distribution in the interval $[0, 1]$. The delay bound for τ_{ca} is $\bar{d} = 2$ to retain the controllability of the system [3] with a mean value of 12.5 ms as shown in Figure 3 (a). The packet loss γ_{ca} follows the Bernoulli distribution with $p = 0.8$ as shown in Figure 3 (b). The plot in Figure 3 is shown for 2 sec for clarity where 1 indicates the packet is received and 0 indicates the packet is lost. The sensor sampling time was selected to be $T_s = 0.01$ sec. A quadratic cost function (4) was selected with $P_z = 10^{-3} I_{8\times 8}$ and $R_z = 10^{-3} I_{2\times 2}$ where I is the identity matrix. The learning gain α_V was selected as 0.05 satisfies $0 < \alpha_V < 2$, $\Gamma = 0.99$, and $\mu = 0.25$. A Monte Carlo simulation is run for 25 sec or 2500 sampled instants.

The performance of the event sampled optimal regulator is shown in Figures 3 to 6. The state vector is regulated to zero by the proposed regulator as shown in Figure 4 (a). It is clear that this event sampled regulator is able to handle random delays and packet losses in the presence of uncertain system dynamics. The optimal control policy is shown in Figure 4 (b). The convergence of the event sampled Bellman error to zero

shown in Figure 4(c) indicates that the optimal cost is attained with the designed regulator.

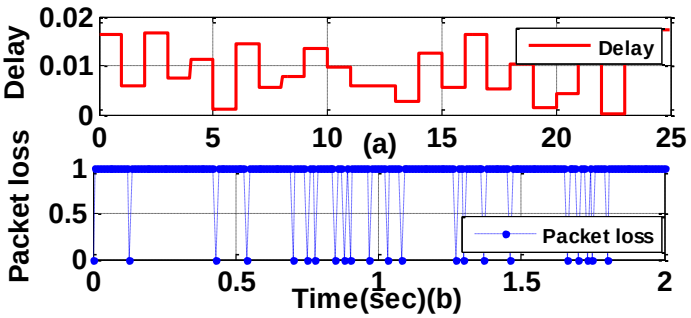

Figure 3. Distribution of (a) delays; and (b) packet losses.

The evolution of the event sampling threshold and error are plotted in Figure 5 (a). The event sampled or the transmission instants are illustrated in Figure 5(b). The vertical lines shows the inter-sample times. The minimum inter-sample-time observed is the sensor sampling time, i.e., 0.01 sec. Further, it is clear that nontrivial inter-sample time exist (height of the vertical line) as discussed in Remark 4. The mean value of the cumulative number of event sampled instants during the simulation time is found to be 853 as shown in Figure 6 (a). Therefore, the transmissions and computations are reduced when compared with the traditional discrete time systems. A comparison of computation using the mean value in terms of the additions and multiplication is shown in Table 1. A 30% reduction of the computation is shown in case of event sampled NCS when compared to its periodic implementation.

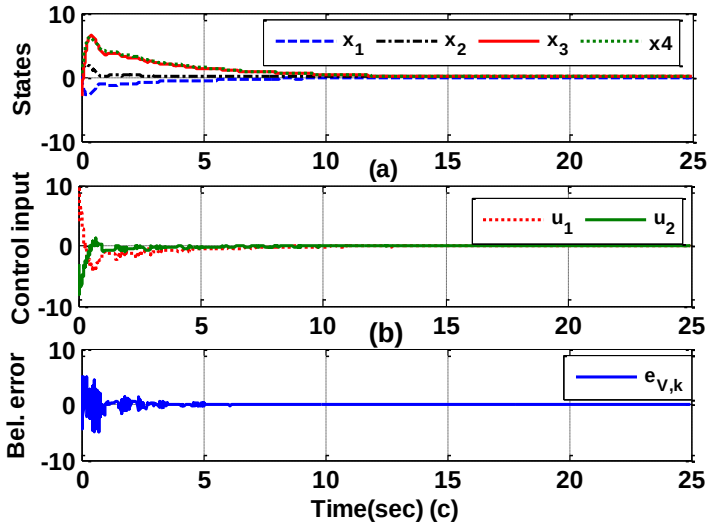

Figure 4. Convergence of: (a) closed-loop state vector; (b) event sampled optimal control policy; and (c) event sampled Bellman error.

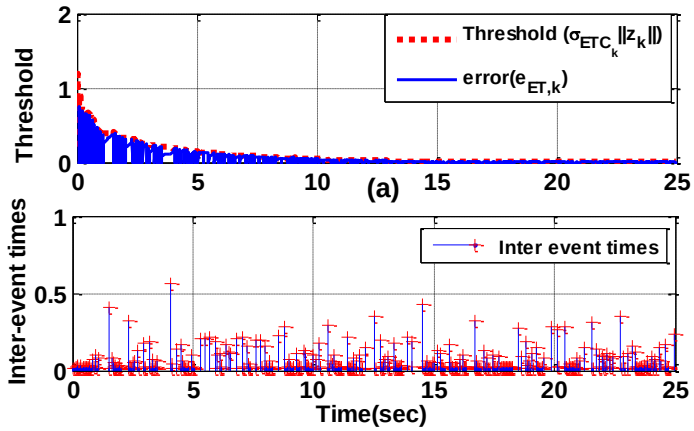

Figure 5. Evolution of (a) the threshold and event sampling error; and (b) inter-sample times.

In terms of bandwidth usage, assuming a packet size of 8 bit data, Figure 6(b) depicts the comparison of data rate in bits per sec (bps) between the event sampled and the traditional periodic schemes. It is evident that the proposed event sampled scheme has a low average data rate. A saving of 56% in the bandwidth usage was observed during the simulation time.

Table 1. Comparison of computational load between traditional periodic sampled and event-based non-periodic sampled system

System		Traditional periodic sampled	Event-based non-periodic sampled
Samping instants		2500	853
Number of additions and Multiplications at every sampling instant	VFE	13	13
	Controller	3	3
	Update law Controller and Mirror	65	65*2
	Trig. Con	0	7
Total number of Computation		202500	142038

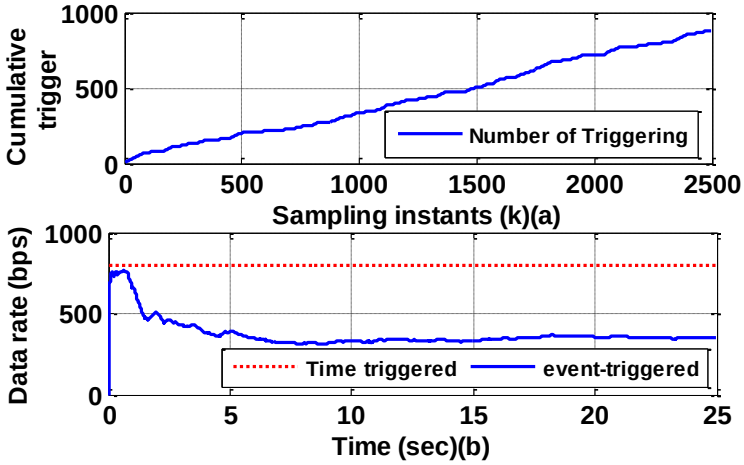

Figure 6. Comparison of the: (a) cumulative number of events with sampled instants; and (b) data rate between periodic and event-triggered system.

5. CONCLUSIONS

In this paper, we presented an event sampled optimal regulator design of an NCS with time varying delays and random packet losses. The novel adaptive event sampled condition along with the tuning law is able to regulate the parameter estimation error and the Bellman error. Finally, the simulation results substantiated the analytical design. It was found that the event sampled Bellman error converged to zero guaranteeing optimal solution. Further, it was determined that the proposed event sampled adaptive optimal regulator is not only able to handle the delays and packet losses but also helps to regulate the state vector while saving the network bandwidth and the computation thereby validating the analytical design.

6. REFERENCES

[1] L. W. Liou and A. Ray, "A stochastic regulator for integrated communication and control systems: part I—formulation of control law," *ASME Journal of Dynamic Systems, Measurement, and Control*, vol. 113, no. 4, pp. 604–611, Jan. 1991.

[2] H. Shousong and Z. Qixin, "Stochastic optimal control and analysis of stability of networked control systems with long delay," *Automatica*, vol. 39, no. 11, pp. 1877-1884, Nov.. 2003.

[3] Hao Xu, S. Jagannathan, and F. L. Lewis, "Stochastic optimal control of unknown linear networked control system in the presence of random delays and packet losses," *Automatica*, vol. 48, no. 6, pp. 1017-1030, Jun. 2012.

[4] Hao Xu and S. Jagannathan, "Stochastic optimal design for unknown linear discrete-time system zero-um games in input-output form under communication constraints," *Asian Journal of Control*, vol. 16, no. 5, pp. 1263-1276, 2014.

[5] Q. Yang and S. Jagannathan, "Reinforcement learning controller design for affine nonlinear discrete-time systems using online approximators," *IEEE Transactions on Systems, Man and Cybernetics: Part B*, vol. 42, no. 2, pp. 377-390, Apr. 2012.

[6] S. Mehraeen, T. Dierks, S. Jagannathan, and Mariesa Crow, "Zero-sum two-player game theoretic formulation of affine nonlinear discrete-time systems using neural networks," *IEEE Transactions on Systems, Man and Cybernetics, Part B* , vol. 43, no. 6, pp. 1641-1655, Dec. 2013.

[7] P. He and S. Jagannathan, "Reinforcement learning-based neural network controller for nonlinear discrete-time systems with input constraints," *IEEE Transactions on Systems, Man and Cybernetics-Part B*, vol. 37, no. 2, pp. 425-437, Apr. 2007.

[8] D. P. Bertsekas and S. E. Shreve, *Stochastic Optimal Control: The Discrete Time Case*. Belmont, MA, USA: Athena Scientific, 1996.

[9] J. Bradtke, B.E. Ydestie, and A.G. Barto, "Adaptive linear quadratic control using policy iteration,' in *Proceeding of the American. Control Conf*erence, Baltimore, MD, USA, Jun.1994, pp. 3475–3476.

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

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

[12] X. Wang and M. D. Lemmon, "On event design in event-triggered feedback systems," *Automatica*, vol. 47, no. 10, pp. 2319-2322, Oct. 2011.

[13] A. Molin and S. Hirche, “On the optimality of certainty equivalence for event-triggered control systems,” *IEEE Transactions on Automatic Control*, vol. 58, no. 2, pp. 470-474, Feb.. 2013.

[14] O. C. Imer and T. Basar, “To measure or to control: optimal control with scheduled measurements and controls,” in *Proceeding of the American. Control Conference,* Minneapolis, MN, USA, Jun. 2006, pp. 14-16.

[15] G. C. Goodwin and K. S. Sin, *Adaptive Filtering Prediction and Control*, 1st edition. Mineola, NY: Dover publications, 1984.

[16] S. Jagannathan, *Neural Network Control of Nonlinear Discrete-Time Systems.* Boca Raton, FL, USA: CRC Press, 2006.

[17] D. Chatterjee and D. Liberzon, “Stability analysis of deterministic and stochastic switched systems via a comparison principle and multiple Lyapunov functions,” *SIAM Journal on Control and Optimization*, vol. 45, no. 1, pp. 174-206, Jul. 2006.

APPENDIX

Proof of Theorem 1: The proof is carried out by considering a single Lyapunov function candidate and evaluating it both at the event sampled instants (with parameter update) and inter-sample times (no update). Both cases are combined together to show the asymptotic convergence of the Q-function parameter estimation error.

Case I: *At the event sampled instants* ($k = k_i$, $\forall i = 1, 2, \cdots$)

Consider the Lyapunov function candidate given as

$$L_{\tilde{\bar{g}},k} = \tilde{\bar{g}}_k^T \tilde{\bar{g}}_k, \; k = k_i, \forall i = 1, 2, \cdots. \tag{A.1}$$

The first difference, $\Delta L_{\tilde{\bar{g}},k} = \underset{\tau,\gamma}{E}\{\tilde{\bar{g}}_{k+1}^T \tilde{\bar{g}}_{k+1}\} - \tilde{\bar{g}}_k^T \tilde{\bar{g}}_k$, along the Q-function parameter estimation error dynamics (22) for the case $k = k_i$ becomes

$$\Delta L_{\tilde{\bar{g}},k} = \underset{\tau,\gamma}{E}\{(\tilde{\bar{g}}_k + (\alpha_V Z_k E_k^{V^T} / \|I + Z_k^T Z_k\|))^T (\tilde{\bar{g}}_k + (\alpha_V Z_k E_k^{V^T} / \|I + Z_k^T Z_k\|))\} - \tilde{\bar{g}}_k^T \tilde{\bar{g}}_k, \; k = k_i$$

Substituting the Q-function error, $E_{V,k}$, from (23), the first difference is given by

$$\Delta L_{\tilde{\bar{g}},k} \le -\alpha_V (2 - \alpha_V) \tilde{\bar{g}}_k^T \underset{\tau,\gamma}{E}\{Z_k Z_k^T / \|I + Z_k^T Z_k\|\} \tilde{\bar{g}}_k, k = k_i .$$

Observe that Z_k satisfies the PE condition. Therefore, it holds that $\bar{Z}_{\min} \le \underset{\tau,\gamma}{E}\{\|Z_k Z_k^T\| / \|I + Z_k^T Z_k\|\} \le 1, \; k = k_i$. Then, the first difference is bounded above by

$$\Delta L_{\tilde{\bar{g}},k} \le -\alpha_V (2 - \alpha_V) \bar{Z}_{\min}^2 \left\|\tilde{\bar{g}}_k\right\|^2 < 0, \; k = k_i , \tag{A.2}$$

by selecting the learning gain $0 < \alpha_V < 2$. By Lyapunov theorem [16], the Q-function parameter estimation error is bounded at the event sampled instants. Next, we will evaluate the same Lyapunov function during the inter-sample times.

Case II: *During inter-sample times* ($k_i < k < k_{i+1}$)

Consider the same Lyapunov function as in Case I. The first difference along the Q-function parameter estimation error dynamics (22) for $k_i < k < k_{i+1}$ is given by

$$\Delta L_{\tilde{\bar{g}},k} = \underset{\tau,\gamma}{E}\{\tilde{\bar{g}}_{k+1}^T \tilde{\bar{g}}_{k+1}\} - \tilde{\bar{g}}_k^T \tilde{\bar{g}}_k = 0,\ k_i < k < k_{i+1}. \tag{A.3}$$

Note that, initial value $\tilde{\bar{g}}_0$ is bounded since $\hat{\bar{g}}_0$ is initialized in a compact set and the target parameters $\bar{g}_k$ are finite.

Next, by combining both the cases we need to show that $\left\|\tilde{\bar{g}}_k\right\| \to 0$ as $k \to \infty$ with trigger of events. From Case I, $\Delta L_{\tilde{\bar{g}},k_i} < 0$, this implies $\left\|\tilde{\bar{g}}_{k_i+1}\right\| < \left\|\tilde{\bar{g}}_{k_i}\right\|$. The first difference (A.2) can be expressed as

$$\Delta L_{\tilde{\bar{g}},k_i} = \underset{\tau,\gamma}{E}\{\tilde{\bar{g}}_{k_i+1}^T \tilde{\bar{g}}_{k_i+1}\} - \tilde{\bar{g}}_{k_i}^T \tilde{\bar{g}}_{k_i} \le \left\|\tilde{\bar{g}}_{k_i+1}\right\|^2 - \left\|\tilde{\bar{g}}_{k_i}\right\|^2 \le -\alpha_V(2-\alpha_V)\bar{Z}_{\min}^2 \left\|\tilde{\bar{g}}_{k_i}\right\|^2.$$

Rearranging the above equation leads to

$$\left\|\tilde{\bar{g}}_{k_i+1}\right\|^2 \le (1-\beta)\left\|\tilde{\bar{g}}_{k_i}\right\|^2,\ k_i < k < k_{i+1}, \tag{A.4}$$

where $0 < \beta = \alpha_V(2-\alpha_V)\bar{Z}_{\min}^2 < 1$. From Case II, $\Delta L_{\tilde{\bar{g}},k} = 0$, $k_i < k < k_{i+1}$, $\forall i = 1,2,\cdots$, therefore, $\left\|\tilde{\bar{g}}_{k_i+1}\right\| = \left\|\tilde{\bar{g}}_k\right\| = \left\|\tilde{\bar{g}}_{k_{i+1}}\right\|, k_i < k < k_{i+1}$. By solving this difference inequality (A.4) recursively with initial condition $\left\|\tilde{\bar{g}}_{k_0}\right\| = \left\|\tilde{\bar{g}}_0\right\| = B_{\tilde{\bar{g}},0}$ and using comparison Lemma [17], (A.4) satisfies

$$\left\|\tilde{\bar{g}}_{k_i+1}\right\|^2 \le \beta\left\|\tilde{\bar{g}}_{k_{i-1}+1}\right\|^2 \le (1-\beta)^{i+1}\left\|\tilde{\bar{g}}_0\right\|^2 \equiv B_{\tilde{\bar{g}},k_i+1},\ \forall i = 1,2,\cdots. \tag{A.5}$$

Observe that $B_{\tilde{\bar{g}},k_i+1}$ is a piecewise constant and converging sequence of functions since $0 < 1-\beta < 1$. This implies $\lim_{k_i \to \infty} B_{\tilde{\bar{g}},k_i+1} = 0$. It follows that $\tilde{\bar{g}}_{k_{i+1}} \to 0$ as the event

sampled instants $k_i \to \infty$. Since k_i is a subsequence of $k \in \mathbb{N}$, $k_i \to \infty$ implies $k \to \infty$. From the above, we can conclude that $\tilde{\tilde{g}}_k \to 0$ as $k \to \infty$. ■

Proof of Theorem 2: To show the asymptotic stability in the mean we will prove that conditions in (26) holds by defining a piecewise continuous upper bound function on the Lyapunov function for all time $k \in \mathbb{N}$.

Case 1: *At the event sampled instants* ($k = k_i, \forall i = 1, 2, \cdots$)

Consider a positive definite Lyapunov function candidate

$$L_k = \Lambda_1 L_{z,k} + \Lambda_2 L_{\tilde{\tilde{g}},k}, \tag{A.6}$$

where $L_{z,k} = z_k^T z_k$ and $L_{\tilde{\tilde{g}},k} = \tilde{\tilde{g}}_k^T \tilde{\tilde{g}}_k$. The positive constant coefficients $\Lambda_1 = \varpi / 2B_{z,\max} C_{\tilde{\tilde{g}}}^2 \Upsilon_M^2 W_M^2$ and $\Lambda_2 = 2\varpi / \alpha_V \left(2 - \alpha_V\right) \bar{Z}_{\min}$ with $0 < \alpha_V < 2$, $0 < \varpi < 1$, $C_{\tilde{\tilde{g}}}$, Υ_M, and W_M are positive constants defined during the proof.

The first difference of the first term $\Delta L_{z,k} = \underset{\tau,\gamma}{E}\{z_{k+1}^T z_{k+1}\} - \Lambda_1 z_k^T z_k$, $k = k_i$ along the closed-loop dynamics (24) for $k = k_i$, i.e., with $\hat{z}_k = z_k$ and $e_{ET,k} = 0$, can be written as

$$\Delta L_{z,k} = \underset{\tau,\gamma}{E}\{(A_{z,k} z_k + B_{z,k} u_{p,k}^* + B_{z,k} \tilde{K}_k z_k)^T (A_{z,k} + B_{z,k} u_{p,k}^* + B_{z,k} \tilde{K}_k z_k)\} - z_k^T z_k,\ k = k_i.$$

By applying Cauchy-Schwartz (C-S) inequality and using Frobenius norm, the first difference leads to

$$\Delta L_{z,k} \le 2 \underset{\tau,\gamma}{E}\{\left\| A_{z,k} z_k + B_{z,k} u_{p,k}^* \right\|^2\} - \left\| z_k \right\|^2 + 2 \underset{\tau,\gamma}{E}\{\left\| B_{z,k} \tilde{K}_k z_k \right\|^2\}$$

$$\le -\Lambda_1 (1 - 2\mu) \left\| z_k \right\|^2 + 2\Lambda_1 B_{z,\max}^2 \left\| \tilde{K}_k z_k \right\|^2,\ k = k_i, \tag{A.7}$$

where $\left\| B_{z,k} \right\| \le B_{z,\max}$ and $\underset{\tau,\gamma}{E}\{\left\| A_{z,k} z_k + B_{z,k} u_{p,k}^* \right\|^2\} \le \mu \left\| z_k \right\|^2$.

The first difference of the second term, $\Delta L_{\tilde{\bar{g}},k}$, remains same as in (A.2) of Theorem 1.

Now, combining (A.7) and (A.2) the overall first difference is given by

$$\Delta L_k \le -\Lambda_1(1-2\mu)\|z_k\|^2 + 2\Lambda_1 B_{z,\max}^2 \left\|\tilde{K}_k z_k\right\|^2 - \Lambda_2 \alpha_V (2-\alpha_V)\bar{Z}_{\min}\left\|\tilde{\bar{g}}_k\right\|^2, \quad k = k_i. \quad \text{(A.8)}$$

Since $\tilde{K}_k$ is a function of Q-function parameter estimation error $\tilde{\bar{g}}_k$, it holds that $\left\|\tilde{K}_k\right\| = \left\| f\left(\tilde{\bar{g}}_k\right)\right\| \le C_{\tilde{\bar{g}}}\left\|\tilde{\bar{g}}_k\right\|$ where $C_{\tilde{\bar{g}}} > 0$ by Lipschitz continuity. Further, the initial control input u_0 is admissible, thus, the time history $\left\|\begin{bmatrix} A_{z,k-1} & B_{z,k-1}\end{bmatrix}\right\| \le \Upsilon_M$ and $\|w_{k-1}\| \le W_M$ where $w_{k-1} = [z_{k-1}^T \quad u_{k-1}^T]$. By using the above facts and substituting the definition of Λ_1 and Λ_2 from (A.6), (A.15) is bounded as

$$\begin{aligned}\Delta L_k &\le -\Lambda_1\left(1-2\mu\right)\|z_k\|^2 + 2\Lambda_1 B_{z,\max}^2 C_{\tilde{\bar{g}}}^2 \Upsilon_M^2 W_M^2 \left\|\tilde{\bar{g}}_k\right\|^2 - \Lambda_2 \alpha_V^s (2-\alpha_V^s)\bar{Z}_{\min}\left\|\tilde{\bar{g}}_k\right\|^2 \\ &\le -\Lambda_1\left(1-2\mu\right)\|z_k\|^2 - \varpi\left\|\tilde{\bar{g}}_k\right\|^2.\end{aligned} \quad \text{(A.9)}$$

for $k = k_i$ where $0 < \varpi < 1$ is a constant defined in (A.6). Since $0 < \mu < 1/3$, it is evident from (A.9) that $\Delta L_k < 0$ and, hence, the Lyapunov function L_k, $k = k_i$ is a non-increasing function, i.e., $L_{k_i+1} < L_{k_i}$.

Case 2: *During the inter-sample times* ($k_i < k < k_{i+1}$)

Consider the same Lyapunov function as in Case I. The first difference of the first term along (24) for $k_i < k < k_{i+1}$ and with simple mathematical manipulation using C-S inequality and Frobenius norm is upper bounded by

$$\Delta L_k \le -(1-3\mu)\|z_k\|^2 + 3\underset{\tau,\gamma}{E}\{\left\|B_{z,k}\tilde{K}_k\right\|^2\}\|z_k\|^2 + 3\Lambda_1 B_{z,\max}^2 \left\|\hat{K}_k\right\|^2 \left\|e_{ET,k}\right\|^2.$$

Recalling the event sampling condition (29) and substituting in the above first difference, one can further obtain

$$\Delta L_k \le -\Lambda_1(1-\Gamma)(1-3\mu)\|z_k\|^2 + 3\Lambda_1 B_{z,\max}^2 C_{\tilde{\bar{g}}}^2 \Upsilon_M^2 W_M^2 \left\|\tilde{\bar{g}}_k\right\|^2,$$

for $k_i < k < k_{i+1}$. From (A.5) in Theorem 1, we have $\left\|\tilde{\bar{g}}_k\right\| = \left\|\tilde{\bar{g}}_{k_i+1}\right\| \le B_{\tilde{\bar{g}},k_i+1}$, $k_i < k < k_{i+1}$. Then, the first difference is bounded above by

$$\Delta L_k \le -\Lambda_1\left(1-\Gamma\right)(1-3\mu)\|z_k\|^2 + B_{\tilde{\bar{g}},k_i+1}^{cl}, k_i < k < k_{i+1}, \tag{A.10}$$

where $B_{\tilde{\bar{g}},k_i+1}^{cl} = 3\Lambda_1 B_{z,\max}^2 C_{\tilde{\bar{g}}}^2 \Upsilon_M^2 W_M^2 B_{\tilde{\bar{g}},k_i+1}$.

Next, the first difference of the second term is same as in (A.3) of Theorem 1.

Finally, the overall first difference, by combining the (A.18) and (A.3), can be expressed as

$$\Delta L_k \le -\Lambda_1\left(1-\Gamma\right)(1-3\mu)\|z_k\|^2 + B_{\tilde{\bar{g}},k_i+1}^{cl}, k_i < k < k_{i+1}. \tag{A.11}$$

From (A.11), the first difference of the Lyapunov function $\Delta L_k < 0$, as long as $\|z_k\| > \sqrt{B_{\tilde{\bar{g}},k_i+1}^{cl} / \Lambda_1(1-\Gamma)(1-3\mu)} \equiv B_{k_i+1}^M$, $k_i < k < k_{i+1}$. By Lyapunov theorem [16], z_k and $\tilde{\bar{g}}_k$, $k_i < k < k_{i+1}$ are bounded. Observe that, $B_{\tilde{\bar{g}},k_i+1}$ in (A.5), $B_{\tilde{\bar{g}},k_i+1}^{cl}$ and $B_{k_i+1}^M$ in (A.11) are piecewise constant for i^{th} inter-sample time. Again, from (A.11), it follows that, $\|z_k\|$ outside the ball of radius $B_{k_i+1}^M$, $k_i < k < k_{i+1}$ will decrease to the ball. Therefore, the Lyapunov function L_k, $k_i < k < k_{i+1}$, $\forall i = 1, 2, \cdots$ will converge to a bound $B_{L,k}$ in finite time where $B_{L,k}$ is define as

$$B_{L,k} = \Lambda_1 (B_{k_i+1}^M)^2 + \Lambda_2 B_{\tilde{g},k_i+1}, k_i < k < k_{i+1}. \tag{A.12}$$

The bound $B_{L,k}$ in (A.12) is computed from (A.6) using the upper bound $B_{k_i+1}^M$ of z_k from (A.11) and $B_{\tilde{g},k_i+1}$ of $\tilde{\tilde{g}}_k$ from (A.5) for $k_i < k < k_{i+1}$.

From the above, we can define a piecewise continuous function h_k such that the conditions (31) holds. Consider the Case I and Case II and define a piecewise continuous function

$$h_k = \max\left\{L_k, B_{L,k}\right\},\ \forall k \in \mathbb{N}. \tag{A.13}$$

It is clear that h_k is positive definite and $h_k \geq L_k$, $\forall k \in \mathbb{N}$. Further, from Case I, $L_{k_i+1} < L_{k_i}$ and Case II, $L_{k_i+1} \to B_{L,k}$, $k_i < k < k_{i+1}$ in a finite time. Furthermore, from Theorem 1, we have $B_{\tilde{g},k_i+1} \to 0$ as $k_i \to \infty$. This implies, $B_{k_i+1}^M \to 0$ and, hence, $B_{L,k} \to 0$ as $k_i \to \infty$. Therefore, $h_k \to 0$ as $k_i \to \infty$ or, alternatively, $k \to \infty$.

Consequently, from both the cases, the closed-loop system state and Q-function parameter estimation error converge to zero asymptotically in the mean.

Finally to show that, $u_k \to u_{e,k}^*$, consider the difference

$$\left\|u_{e,k}^* - u_k\right\| = \left\|K_k^* \hat{z}_k - \hat{K}_k \hat{z}_k\right\| \leq C_{\tilde{\tilde{g}}} \left\|\tilde{\tilde{g}}_k\right\| \left\|\hat{z}_k\right\|.$$

Since, $\left\|\tilde{\tilde{g}}_k\right\| \to 0$ $u_k \to u_{e,k}^*$ as $k \to \infty$. ■

Proof of Corollary 1: The event sampling condition (27) can be rewritten as

$$\sigma_{ET,k} \left\|e_{ET,k}\right\| + \left\|e_{ET,k}\right\| \leq \sigma_{ET,k} \left\|\hat{z}_k\right\|.$$

By rearranging the expression one can reach at

$$\left\|e_{ET,k}\right\| \leq \sigma_{ET,k} \left(\left\|\hat{z}_k\right\| - \left\|e_{ET,k}\right\|\right) \leq \sigma_{ET,k} \left\|\hat{z}_k + e_{ET,k}\right\| = \sigma_{ET,k} \left\|z_k\right\|$$

Since the hypothesis of Theorem 1 holds, the closed-loop NCS is asymptotically stable in the mean. ■

Proof of Theorem 3: Consider the event sampling error (14). The error dynamics for $k_i < k < k_{i+1}$ is given by

$$\begin{aligned} e_{ET,k+1} &= \underset{\tau,\gamma}{E}\{A_{z,k} z_k + B_{z,k} u^*_{p,k} + B_{z,k}\tilde{K}_k z_k + B_{z,k}\hat{K}_k e_{ET,k} - z_k\} \\ &\le (\sqrt{\mu} + B_{z,\max} K^*_{\max})\left\|e_{ET,k}\right\| + (B_{z,\max}\left\|\tilde{K}_k\right\| + 1 + \mu)\left\|\hat{z}_k\right\| \end{aligned}.$$

where $\left\|K^*_k\right\| \le K^*_{\max}$. Recalling (13) and substituting in the above equation, it reveals that

$$\left\|e_{ET,k+1}\right\| \le F\left\|e_{ET,k}\right\| + M_i\left\|z_{k_i}\right\|,\ k_i < k < k_{i+1}, \tag{A.14}$$

where $F = \sqrt{\mu} + B_{z,\max} K^*_{\max}$ and $M_i = (\sqrt{\mu} + B_{\max}\left\|\tilde{K}_{k_i}\right\| + 1)$. The variables z_{k_i} and M_i in (A.14) are constant during the i^{th} inter-sample time. Hence, (A.14) becomes a stochastic difference equation with constant input. The solution of difference equation, (A.14) by using comparison lemma [17], is bounded above by

$$\left\|e_{ET,k}\right\| \le \sum_{j=k_i}^{k-1} M_i\left\|z_{k_i}\right\|(F)^{k-j-1} = (F^{k-k_i} - 1)M_i\left\|z_{k_i}\right\|/(F-1). \tag{A.15}$$

Now at the k_{i+1} -th event sampled instant we have $\left\|e_{ET,k_{i+1}}\right\| = \sigma_i\left\|z_{k_i}\right\|$ where $\sigma_i = \sigma_{ET,k_i}/(1+\sigma_{ET,k_i})$ is the threshold coefficient, from (27), for i^{th} inter-sample time.

Then, expressing (A.15) for k_{i+1} time instant, it holds that

$$[(F)^{k_{i+1}-k_i} - 1]M_i\left\|z_{k_i}\right\|/(F-1) \ge \sigma_i\left\|z_{k_i}\right\|.$$

By solving this, the inter-sample times δk_i satisfies

$$\delta k_i \ge \ln\left(1 + (1/M_i)(F-1)\sigma_i\right)/\ln(F)\ ,\ \forall i = 1, 2, \cdots. \tag{A.16}$$

Further, from (A.16), the inter-sample times becomes non-trivial when

$$\ln\left(1+(1/M_i)(F-1)\sigma_i\right) > \ln(F)$$

Alternatively, $\sigma_i / M_i > 1$. ■

APPENDIX B

OPTIMAL REGULATION OF NONLINEAR NETWORKED CONTROL SYSTEMS BY USING EVENT DRIVEN ADAPTIVE DYNAMIC PROGRAMMING

B. OPTIMAL REGULATION OF NONLINEAR NETWORKED CONTROL SYSTEMS BY USING EVENT-DRIVEN ADAPTIVE DYNAMIC PROGRAMMING

Avimanyu Sahoo and S. Jagannathan

Abstract *— In this paper a stochastic near optimal regulation of nonlinear networked control systems (NCS) is presented with event sampled state and input vector. Event-driven stochastic adaptive dynamic programming (ADP) based technique is utilized with neural networks (NN) to design the near optimal policy. An actor-critic framework with event sampled feedback information is utilized to implement the ADP scheme. The system dynamics are approximated by using a novel NN identifier with event sampled inputs. The stochastic identifier, actor, and critic NN weights are tuned at the event-sampled instants leading to aperiodic tuning laws. Above all, an adaptive event sampling condition based on estimated NN weights is designed by using the Lyapunov technique to ensure ultimate boundedness of all the closed-loop signals and to ensure approximation accuracy. The net result is event-driven approximate dynamic programming technique that can significantly reduce the computation and network transmissions. Finally, the analytical design is substantiated with simulation results.*

Index Terms **-** Adaptive dynamic programming, event sampled control, neural networks, optimal control.

1. INTRODUCTION

The presence of the packet switched communication network between the plant and the controller in a networked control system (NCS) [1]-[7] brings in the unavoidable network constraints. These constraints are of the form of time-varying network induced delays, random packet losses, quantization error and congestion. A traditional controller may jeopardize the stability and performance of the NCS in the presence of these artifacts. The effect of these constraints on stability of the linear and nonlinear systems is studied in detail by various authors [1]-[7].

Traditional optimal control methods are also extended to NCS [3] by using stochastic Riccati equation. Since, the backward-in-time solution is not preferred in practical implementation, adaptive dynamic programming (ADP) [5]-[6], [8]-[15] in conjunction with reinforcement learning technique is used for a forward-in-time and online solution. In general, policy and value iteration based techniques [9] are used to solve the Hamilton-Jacobi-Bellman (HJB) equation to compute the optimal control policy. These iterative approaches are computational intensive and may require the complete knowledge of the system dynamics.

In the recent years, the ADP scheme is also extended to the NCS with time-varying network induced delays and random packet losses where the system dynamics become uncertain and stochastic [3]-[6]. A time-driven ADP scheme without using the knowledge of the system dynamics and policy/value iteration is presented in [4]-[6] both for linear [4] and nonlinear systems [5]-[6]. Although these schemes [3]-[6] result in satisfactory performance and stability requirement under periodic sampling and

transmission, they, however, lead to higher computational cost and require considerable network bandwidth.

State based sampling and transmission scheme when used for control, formally known as event-triggered control (ETC) [16]-[21], is found to be effective in terms of resource utilization with a certain level of performance guarantee. The ETC scheme for NCS with inherent network constraints such as constant or time-varying delays and packet losses presented in [16]. A reduction in network bandwidth usages with asymptotic stability for bounded delays and packet losses is shown. The design assumed the complete knowledge of the system dynamics for implementation. On the other hand, optimal control with event sampled transmission is studied by various authors [19]-[22]. The traditional backward-in-time solution of the Riccati equation (RE) is used under the assumption of separation principle [20], [22]. To the best knowledge of the authors, no known event-driven ADP scheme is available for nonlinear NCS (NNCS).

Therefore, in this paper, we propose a stochastic event sampled optimal regulator design for NNCS with network induced time-varying delays and random packet losses. Stochastic actor-critic neural network (NN) based ADP scheme is introduced with event sampled state and control input vector. The optimal regulator is designed for systems with completely uncertain system dynamics and time delays greater than one sensor sampling instant in the network between the controller and actuator. The main differences between the ADP scheme presented in this paper and that in [5]-[6] include: the NN-based approximation of the dynamics with event sampled state and input vectors and the NN weight tuning only at the event sampling instants leading to an aperiodic update. Therefore, the ADP scheme in this paper requires the design of a novel event sampling

criterion in order to facilitate the event sampled NN approximation and retain the advantage of the traditional ETC [16]-[18].

First, the universal NN approximation property is revisited in the context of event-based sampling. The system dynamics, the control policy, and the value function are approximated using the event sampled identifier, actor, and critic NNs, respectively. The NN weights are tuned at the sampled instants to force the Hamilton-Jacobi-Bellman (HJB) error to a minimum. The adaptive event sampling condition is designed based on the approximation accuracy and the system stability to determine the sampling instants. This adaptive sampling condition assists the approximation with reduced communication and computation. Finally, the ultimate boundedness (UB) in the mean of the closed-loop event sampled system is presented using the Lyapunov technique. A preliminary version of the work is published in [23].

The remaining of the paper is organized as follows. Section 2 presents a brief background on stochastic ADP and formulates the problem. Section 3 presents the detailed design procedure. Before concluding in Section 5, simulation results are included in Section 4. The appendix gives details of the proof of lemmas and theorems.

2. BACKGROUND AND PROBLEM FORMULATION

In this section, a brief background on the reformulation of the NCS dynamics incorporating the delays and packet losses is presented. Then, the problem we considered in the paper is formulated.

2.1 NNCS REFORMULATION

Consider the NNCS represented by a nonlinear continuous-time system described by

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

where $x(t) \in \Re^n$ and $u(t) \in \Re^m$ represent the system state and the control input vectors, respectively. The nonlinear functions $f(x(t)) \in \Re^n$ and $g(x(t)) \in \Re^{n\times m}$ denote the internal dynamics and the control coefficient function, respectively, with $f(0) = 0$ at $x = 0$ being the unique equilibrium point. It is considered that the functions $f(x(t))$ and $g(x(t))$ are unknown with the following standard assumption.

Assumption 1[6]: The system (1) is controllable and observable and the system state vector is measureable. The matrix $g(x)$ satisfies $\|g(x)\| \le g_M$ in a compact set for all $x \in \Omega_x$ with $g_M > 0$ is a known constant.

Considering the networked induced time-varying delays from the sensors to the controller, $\tau_{sc}(t)$, the controller to the actuator, $\tau_{ca}(t)$, and the random packet losses $\gamma(t)$ between controller and actuator, the system in (1) can be expressed as [2]

$$\dot{x}(t) = f(x(t)) + \gamma(t)g(x(t))u(t - \tau_{ca}(t)), \tag{2}$$

The packet losses $\gamma(t)$ is denoted by

$$\gamma(t)=\begin{cases} I^{n\times n}, \text{control input received at actuator at time } t, \\ 0^{n\times n}, \text{control imput lost at time } t. \end{cases}$$

For reformulating the dynamics of the system by incorporating the network constraints, the following properties of the communication network are assumed [2]-[4].

Assumption 2:

(a). The sensor is assumed to be time driven and samples the system state at a fixed sampling interval T_s [2] whereas the feedback and control input is transmitted at event sampled instants;

(b) The communication network used is a wide area network so that the sensor to controller and controller to actuator network induced delays are considered independent and satisfy the following criteria. The delays $\tau_{sc} \le \Delta_s$ [2] and $\tau_{ca} \le \bar{d}T_s$ where $\Delta_s \le T_s$ between the sensor and controller sampling instant is the fixed skew and $\bar{d}$ is a positive integer [2]-[4];

(c) The packet losses between sensor and controller is negligible and the distribution between controller to actuator is known;

(d) The initial system state is deterministic [3][4].

Since, the communication network uses packet switched transmission and for implementation the regulator in a digital platform, a discrete-time formulation of the system is necessary for the control design. The system dynamics (2) can be discretized by integrating it within a sensor sampling interval $[kT_s \;\; (k+1)T_s]$ for all $k \in \mathbb{N}$. The discrete-time representation is given by [6]

$$x_{k+1} = \bar{F}_{\tau,\gamma}(x_k, u_{k-1}, \cdots u_{k-\bar{d}}) + \bar{G}_{\tau,\gamma}(x_k, u_{k-1}, \cdots u_{k-\bar{d}})u_k \,, \tag{3}$$

where $\bar{F}_{\tau,\gamma}(x_k, u_{k-1}, \cdots u_{k-\bar{d}})$ and $\bar{G}_{\tau,\gamma}(x_k, u_{k-1}, \cdots, u_{k-\bar{d}})$ are the transformed NCS dynamics given in [5]-[6]. The vector $x_k = x(kT_s)$ is the discretized state vector, $u((k-l)T_s) = u_{k-l}$, $l = 0,1,\cdots,\bar{d}$ is the delayed control input. It is possible that the control input vectors arrive simultaneously at the actuator. Thus, a packet reordering mechanism selects the latest control input for use.

By selecting an augmented state vector $z_k = [x_k^T \quad u_{k-1}^T \quad \cdots \quad u_{k-\bar{d}}^T]^T$, the discrete-time system (3) can be presented in a compact form as

$$z_{k+1} = F(z_k) + G(z_k)u_k, \tag{4}$$

where $F(z_k) = [\bar{F}_{\tau,\gamma}^T(x_k, u_{k-1}, \cdots, u_{k-\bar{d}}) \; 0^T \; u_{k-1}^T \; u_{k-\bar{d}+1}^T]^T \in \Re^{\bar{d}m+n}$ and $G(z_k) = [\bar{G}_{\tau,\gamma}(x_k, u_{k-1}, \cdots, u_{k-\bar{d}}) \mid I_m \mid 0 \mid \cdots \mid 0] \in \Re^{(\bar{d}m+n)\times m}$ with $[A \mid B \mid C]$ denotes the vertical concatenation of the matrices A, B and C. The matrix I is the identity matrix and ‘ 0 ’ represents the null vectors or the matrices of appropriate dimensions. The transformed system is a stochastic uncertain nonlinear system due to the presence of random packet losses and delays. From Assumption 1 and definition (4) the stochastic matrix function $G(z_k)$ satisfies $\|G(z_k)\| \le G_M$ with $G_M > 0$ is a computable constant [6].

To design an optimal controller for the system in (1) in the presence of the networked induced time-varying delays and packet losses, it is sufficient to design an optimal controller for the augmented NCS dynamics in (4). With this respect, consider the value function given by [3][4]-[6] as

$$V_k = \underset{\tau,\gamma}{E}\left\{\sum_{j=k}^{\infty} z_j^T Q_z z_j + u_j^T R_z u_j\right\}, \tag{5}$$

where $Q_z \in \Re^{n+\bar{d}m \times n+\bar{d}m}$ and $R_z \in \Re^{m \times m}$ are, respectively, symmetric positive semi-definite and definite penalty matrices. The operator $\underset{\tau,\gamma}{E}(\bullet)$ is the expectation operator over all the random delays and packet losses. The initial control, u_0, is assumed to be admissible [15] to keep the value function finite.

The value function (5) can be written as

$$V_k = \underset{\tau,\gamma}{E}\left\{\sum_{j=k}^{\infty} z_j^T Q_z z_j + u_j^T R_z u_j\right\} + V_{k+1} \tag{6}$$

where $V_{k+1} = \underset{\tau,\gamma}{E}\left\{\sum_{j=k+1}^{\infty} z_j^T Q_z z_j + u_j^T R_z u_j\right\}$ is the cost from $k+1$ onwards. The optimal control input by differentiating the optimal value function $V_k^* = \min_{u_k}(\underset{\tau,\gamma}{E}\{z_k^T Q_z z_k + u_k^T R_z u_k\}) + V_{k+1}^*$ can be represented as

$$u_k^* = -(1/2)\underset{\tau,\gamma}{E}\{R_z^{-1} G^T(z_k)\partial V_{k+1}^*/\partial z_{k+1}\}\ . \tag{7}$$

By using (7), the discrete time HJB equation can be expressed as

$$V_k^* = \underset{\tau,\gamma}{E}\{z_k^T Q_z z_k + (1/4)(\partial V_{k+1}^{*T}/\partial z_{k+1}) G(z_k) R_z^{-1} G^T(z_k)(\partial V_{k+1}^*/\partial z_{k+1})\} + V_{k+1}^*. \tag{8}$$

Since, a closed form solution of (8) is quite difficult, stochastic ADP based techniques [5]-[6] are used for an online and forward-in-time solution. Next, the problem of event sampled stochastic ADP is formulated.

2.2 PROBLEM STATEMENT

Our objective is regulate the augmented system dynamics of the NNCS (4) with event sampled feedback information minimizing the value function (5). The event sampled instants can be characterized as a subsequence $\{k_i\}_{i=1}^{\infty}$ of periodic sensor sampling instants $k \in \mathbb{N}$. We will assume the initial event sampling instant is at time

$k_0 = 0$ and the initial state and control input are transmitted. The event sampled state vector denoted by $\breve{x}_k$ held the controller by the zero-order-hold is given by

$$\breve{x}_k = x_{k_i},\ k_i \le k < k_{i+1},\ \ \forall i = 1, 2, \cdots. \tag{9}$$

where $x_{k_i}, \forall i = 1, 2, \cdots$ is the state vector at the event sampled instants..

The event sampled augmented system state vector by using $\breve{x}_k$ can be expressed as

$$\breve{z}_k = z_{k_i},\ k_i \le k < k_{i+1},, \tag{10}$$

where $\breve{z}_k = [\breve{x}_k^T\ u_{k-1}^T \cdots u_{k-\bar{d}}^T]^T$ and $z_{k_i} = [x_{k_i}^T\ u_{k_i-1}^T\ u_{k_i-\bar{d}}^T]^T$. The corresponding error between z_k and $\breve{z}_k$ can be expressed as

$$e_{ET,k} = z_k - \breve{z}_k,\ k_i \le k < k_{i+1}, \forall i = 1, 2, \cdots \tag{11}$$

where $e_{ET,k}$ is referred to as event sampling error . This event sampling error (11) is reset to zero along with the update of $\breve{x}_k$ at the sampling instants and denoted by

$$e_{ET,k} = 0,\ k = k_i, \forall i = 1, 2, \cdots. \tag{12}$$

Since, the system dynamics, the value function, and the control input are unknown, the NN approximation property with event sampled information is revisited next.

Consider a continuous stochastic function $h(z_k) \in \Re^n$ in a compact set for all $z_k \in \Omega_z$. By universal approximation property, the function $h(z_k) \in \Re^n$ in the compact set can be expressed as

$$h(z_k) = \underset{\gamma,\tau}{E}\{W_h^T \upsilon_h(z_k) + \varepsilon_h(z_k)\}, \tag{13}$$

where $\underset{\tau,\gamma}{E}\{W_h\} \in \Re^{l_h \times n}$ is the unknown constant target NN weights, $\upsilon_h(z_k) \in \Re^{l_h}$ is the periodic sampled activation function, and $\varepsilon_h(z_k) \in \Re^m$ is the traditional reconstruction error with l_h as the number of neurons. With event sampled transmission the available state vector $\breve{z}_k$ will be used for approximation. The NN approximation (13) is given by

$$h(z_k) = \underset{\gamma,\tau}{E}\{W_h^T \upsilon_h(\breve{z}_k) + \varepsilon_{e,h}(\breve{z}_k, e_{ET,k})\},\ k_i \le k < k_{i+1} \tag{14}$$

where $\varepsilon_{e,h}(\breve{z}_k, e_{ET,k}) = \underset{\gamma,\tau}{E}\{W_h^T(\upsilon_h(\breve{z}_k + e_{ET,k}) - \upsilon_h(\breve{z}_k)) + \varepsilon_h(z_k)\}$. The error $\varepsilon_{e,h}(\breve{z}_k, e_{ET,k})$ is clearly a function of the event sampling error $e_{ET,k}$ due to the relation $(\upsilon_h(z_k) - \upsilon_h(\breve{z}_k))$ $= (\upsilon_h(\breve{z}_k + e_{ET,k}) - \upsilon_h(\breve{z}_k))$ along with the traditional reconstruction error expressed as $\varepsilon_h(z_k) = \varepsilon_h(\breve{z}_k + e_{ET,k})$. This implies the event sampling error drives the accuracy of event sampled approximation.

By using (14), the event sampled approximation of the value function (5) can be expressed as

$$V_k = \underset{\tau,\gamma}{E}\{W_V^T \phi_V(\breve{z}_k) + \varepsilon_{e,V}(\breve{z}_k, e_{ET,k})\},\ k_i \le k < k_{i+1}, \tag{15}$$

where $\underset{\tau,\gamma}{E}\{W_V\} \in \Re^{l_V}$ is the unknown constant target NN weights, $\underset{\tau,\gamma}{E}\{\phi_V(\breve{z}_k)\} \in \Re^{l_V}$ is the event sampled activation function. The error $\varepsilon_{e,V}(\breve{z}_k, e_{ET,k}) = \underset{\gamma,\tau}{E}\{W_V^T(\phi_V(\breve{z}_k + e_{ET,k}) - \phi_V(\breve{z}_k))$ $+\varepsilon_V(\breve{z}_k + e_{ET,k})\}$ is the event sampled reconstruction error where $\varepsilon_V(z_k) \in \Re$ is the traditional reconstruction error with l_V being the number of neurons.

Substituting the approximated value function (15) into (6), equation (6) can be expressed as

$$\underset{\tau,\gamma}{E}\{W_V^T\phi_V(\breve{z}_k)+\varepsilon_{e,V}(\breve{z}_k,e_{ET,k})\}=\underset{\tau,\gamma}{E}\{z_k^TQ_zz_k+u_k^TR_zu_k\}+\underset{\tau,\gamma}{E}\{W_V^T\phi_V(\breve{z}_{k+1})+\varepsilon_{e,V}(\breve{z}_{k+1},e_{ET,k+1})\}$$

Rewriting the above expression

$$\underset{\tau,\gamma}{E}\{\Delta\varepsilon_{e,V}(\breve{z}_k,e_{ET,k})\}=\underset{\tau,\gamma}{E}\{z_k^TQ_zz_k+u_k^TR_zu_k\}+\underset{\tau,\gamma}{E}\{W_V^T\Delta\phi_V(\breve{z}_{k+1})\},\tag{16}$$

where $\Delta\phi_V(\breve{z}_k)=\phi_V(\breve{z}_{k+1})-\phi_V(\breve{z}_k)$ and $\Delta\varepsilon_{e,V}(\breve{z}_k,e_{ET,k})=\varepsilon_{e,V}(\breve{z}_{k+1},e_{ET,k+1})-\varepsilon_{e,V}(\breve{z}_k,e_{ET,k})$.

One can notice that ideal HJB equation (16) results in a higher reconstruction error $\underset{\tau,\gamma}{E}\{\Delta\varepsilon_{e,V}(\breve{z}_k,e_{ET,k})\}$ due to the event sampling error. From the above discussion, both the accuracy of the approximation and optimality depends upon the event sampling error or in turn the event sampling condition.

In addition, from the NN weight estimation point of view, the NN weights can only be tuned at the event sampled instants since the feedback information is available only at these instants. This further leads to an aperiodic update in contrast to the periodic ones used in traditional NN based approach. Moreover, the ADP scheme requires the control coefficient matrix $G(\mathrm{z}_k)$ to implement the optimal control input (7). Since, the matrix $G(\mathrm{z}_k)$ is unknown there is requirement for identifying the system dynamics also.

From the above discussion, the optimal control problem can be precisely defined as: (a) design a novel event sampling condition which will ensure the accuracy in function approximation, (b) design an identifier to to approximate the control coefficient function $G(z_k)$, and (c) retain the advantages of traditional ETC while ensuring optimality. A solution to the above problem is presented next.

3. EVENT SAMPLED OPTIMAL REGULATOR DESIGN

In this section, a solution for event-based ADP with the detailed design procedure is presented.

3.1 PROPOSED SOLUTION

The structure of the proposed stochastic NN based ADP scheme is illustrated in Figure 1. Event sampled actor-critic NNs architecture will be used for control input and value function approximation, respectively. A novel event sampled NN based identifier is also proposed to approximate the unknown system dynamics.

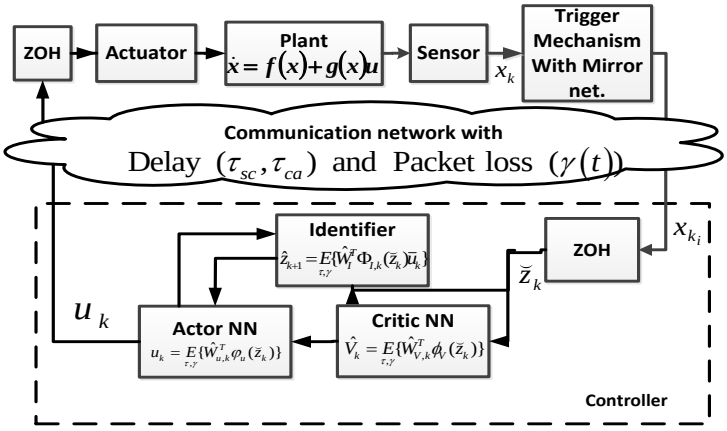

Figure 1. Proposed event-triggered NCS architecture.

For ensuring the accuracy of NN approximation, an adaptive event sampling condition is proposed. The event sampling condition is a function of NN weight estimates and the system state vector, and gets updated with every update of the NN estimates at the event sampled instants. This implies that the event sampling condition is implicitly adjusted by the approximation errors. A detailed discussion is included in Remark 3.

The trigger mechanism at the sensor evaluates the event sampling condition at every periodic sampling instant. A transmission decision is made once the event sampling condition is violated. Since the condition is a function of the actor-critic NN weight estimates, the weights must be transmitted from the controller to the trigger mechanism. To avoid this transmission of the NN weights, a mirror actor-critic NN is used at the trigger-mechanism to estimate the NN weights locally. This further requires synchronization of both the actor-critic NNs to ensure same information at both trigger-mechanism and controller.

Since, the delay from the senor to controller, $\tau_{sc} \leq \Delta_s$, the state information at the sensor's sampling instant and controller's sampling instants are same [2]. Further, by initializing the weights at both the actor-critic networks with same value and updating at event sampled instants synchronism can be achieved. Although, the use of additional mirror actor-critic NN at the trigger mechanism increases the computation when compared to the traditional ADP schemes, the event sampled execution reduces the overall computation.

3.2 EVENT SAMPLED ADP BASED OPTIMAL CONTROL DESIGN

The event sampled stochastic ADP design entails four design steps: identifier, critic, actor, and event sampling condition. The identifier design is presented next.

3.2.1 Event Sampled Identifier Design. The main objective of the identifier design is to approximate the control coefficient matrix function $G(z_k)$ for computing the optimal control policy. Consider the augmented stochastic dynamics (4). By using event sampled NN approximation in (14), the stochastic dynamics of the augmented system can be written as

$$z_{k+1} = F(z_k) + G(z_k)u_k = [F(z_k) \quad G(z_k)][1 \quad u_k^T]^T$$

$$= \underset{\tau,\gamma}{E}\{W_I^T \Phi_I(\breve{z}_k)\bar{u}_k + \bar{\varepsilon}_{e,I}(\breve{z}_k, e_{ET,k})\},\ k_i \le k < k_{i+1}, \tag{17}$$

where $W_I = [W_F^T \quad W_G^T]^T \in \Re^{(m+1)l_I \times (\bar{d}m+n)}$ is the constant target NN matrix with $W_F \in \Re^{l_I \times (\bar{d}m+n)}$ and $W_G \in \Re^{ml_I \times (\bar{d}m+n)}$ are the target for respective functions $F(\bullet)$ and $G(\bullet)$. The matrix function $\Phi_I(\breve{z}_k) = diag\left(\phi_F(\breve{z}_k) \quad \phi_G(\breve{z}_k)\right) \in \Re^{(m+1)l_I \times m}$ is the event sampled activation function with $\phi_F(\breve{z}_k) \in \Re^{l_I}$, $\phi_G(\breve{z}_k) \in \Re^{ml_I \times m}$, and l_I is the number of neurons in the network. The vector $\bar{u}_k = [1 \quad u_k^T]^T \in \Re^{m+1}$ is the augmented control input and $\bar{\varepsilon}_{e,I}(\breve{z}_k, e_{ET,k}) = W_I^T[\Phi_I(\breve{z}_k + e_{ET,k}) - \Phi_I(\breve{z}_k)]\bar{u} + \bar{\varepsilon}_I(\breve{z}_k + e_{ET,k})$ is the event sampled reconstruction error where $\bar{\varepsilon}_I(\breve{z}_k + e_{ET,k}) = \varepsilon_I \bar{u}_k$ with $\varepsilon_I = [\varepsilon_F(z_k)\ \varepsilon_G(z_k)]$ is the traditional reconstruction error. The following assumption holds for the identifier NN in a compact set.

Assumption 3 [24]: The NN identifier target weight matrix, activation function and the traditional reconstruction error are upper bounded in the mean. Then, it holds that $\left\| \underset{\tau,\gamma}{E}\{W_I\} \right\| \le W_{I,M}$, $\left\| \underset{\tau,\gamma}{E}\{\Phi_I(\bullet)\} \right\| \le \Phi_{I,M}$, and $\left\| \underset{\tau,\gamma}{E}\{\varepsilon_I(\bullet)\} \right\| \le \varepsilon_{I,M}$ where $W_{I,M}$, $\Phi_{I,M}$, and $\varepsilon_{I,M}$ are positive constants. In addition, the activation function is Lipschitz continuous in the compact set for all $z_k \in \Omega_z$. Then, there exists a constant C_{Φ_I} such that $\left\| \underset{\tau,\gamma}{E}\{\Phi_I(z_k)\} - \underset{\tau,\gamma}{E}\{\Phi_I(\breve{z}_k)\} \right\| \le C_{\Phi_I} \left\| \underset{\tau,\gamma}{E}\{z_k\} - \underset{\tau,\gamma}{E}\{\breve{z}_k\} \right\|$.

The Lipschitz continuity assumption is satisfied by all NN activation functions and will be used later during the closed-loop stability proof. The event sampled identifier dynamics for the system dynamics (17) can be defined as

$$\hat{z}_{k+1} = \hat{F}(\breve{z}_k) + \hat{G}(\breve{z}_k)u_k, k_i \le k < k_{i+1}. \tag{18}$$

where $\hat{z}_k$ is the identifier state vector, $\hat{F}(\breve{z}_k)$ and $\hat{G}(\breve{z}_k)$ are the identifier dynamics. The function $\hat{F}(\breve{z}_k)$ and $\hat{G}(\breve{z}_k)$ uses the event sampled state vector $\breve{z}_k$ instead of the identifier state $\hat{z}_k$. This architecture reduces unnecessary computation due to the identifier during the inter-sample times.

The identifier with NN weight estimates is given by

$$\hat{z}_{k+1} = \underset{\tau,\gamma}{E}\{\hat{W}_I^T \Phi_{I,k}(\breve{z}_k)\bar{u}_k\},\ k_i \le k < k_{i+1}. \tag{19}$$

Defining the identification error as $\underset{\tau,\gamma}{E}\{\tilde{z}_k\} = \underset{\tau,\gamma}{E}\{z_k - \hat{z}_k\}$, $k_i \le k < k_{i+1}$, the identification error dynamics can be written as

$$\underset{\tau,\gamma}{E}\{\tilde{z}_{k+1}\} = \underset{\tau,\gamma}{E}\{\tilde{W}_{I,k}^T \Phi_I(z_k)\bar{u}_k + \hat{W}_I^T[\Phi_I(z_k) - \Phi_I(\breve{z}_k)]\bar{u} + \bar{\varepsilon}_I(z_k)\}, \tag{20}$$

for $k_i \le k < k_{i+1}$ where $\underset{\tau,\gamma}{E}\{\tilde{W}_{I,k}\} = \underset{\tau,\gamma}{E}\{W_I - \hat{W}_{I,k}\}$ is defined as the identifier NN weight estimation error.

The identifier NN weights are tuned at the event sampled instants such that the identification error converges close to zero in the mean. To achieve this identifier weight update law using the previous values can be selected as

$$\underset{\tau,\gamma}{E}\{\hat{W}_{I,k}\} = \underset{\tau,\gamma}{E}\left\{\hat{W}_{I,k-1} + \chi_k \frac{\alpha_I \Phi_I(\breve{z}_{k-1})\bar{u}_{k-1}\tilde{z}_{I,k}^T}{(\Phi_I(\breve{z}_{k-1})\bar{u}_{k-1})^T(\Phi_I(\breve{z}_{k-1})\bar{u}_{k-1}) + 1}\right\}, \tag{21}$$

where $\alpha_I > 0$ is the learning gain, χ_k is an indicator function defined as

$$\chi_k = \begin{cases} 1, \text{ transmission received} \\ 0, \text{ no transmission} \end{cases} \tag{22}$$

The indicator function $\chi_k = 1$ at the event sampled instants $k = k_i$, $\forall i = 1, 2, \cdots$ and, hence, the NN weights are updated. During the inter-sample times $k_i < k < k_{i+1}$, $\chi_k = 0$ and the NN weights are held. Thus, the NN weights are updated in an aperiodic manner. This reduces computation when compared to the traditional periodic update [6], [15], [24].

By using (21) with a time step forward and the definition $\underset{\tau,\gamma}{E}\{\tilde{W}_{I,k}\} = \underset{\tau,\gamma}{E}\{W_I - \hat{W}_{I,k}\}$, the identifier NN weight estimation error dynamics can be expressed as

$$\underset{\tau,\gamma}{E}\{\tilde{W}_{I,k+1}\} = \underset{\tau,\gamma}{E}\left\{\tilde{W}_{I,k} - \frac{\chi_k \alpha_I \Phi_I(\breve{z}_k)\bar{u}_k \tilde{z}_{I,k+1}^T}{(\Phi_I(\breve{z}_k)\bar{u}_k)^T(\Phi_I(\breve{z}_k)\bar{u}_k)+1}\right\},\ k_i \le k < k_{i+1}. \tag{23}$$

Remark 1: The regression vector $\Phi_I(\breve{z}_k)\bar{u}_k$ must satisfy the persistency of excitation (PE) condition [24] for the identifier NN weight estimation error to converge to zero in the mean. The PE condition requirement is standard in the adaptive and NN based control literature [15], [24]. For completeness the definition of PE condition is presented next.

Definition 1[15]: A vector $\vartheta(z_k)$ is said to be persistently exciting over an interval if there exists positive constants δ , $\underline{\alpha}$, $\bar{\alpha}$ and $k_d \ge 1$ such that $\underline{\alpha} I \le \sum_{k=k_d}^{k+\delta} \vartheta(z_k)\vartheta^T(z_k) \le \bar{\alpha} I$ where I is the identity matrix of appropriate dimension.

The ultimate boundedness (UB) in the mean of the NN weight estimation error is claimed in the following theorem. Before stating the theorem, the following stability notion is necessary.

Definition 2 [5]: An equilibrium point x_e is said to be ultimately bounded (UB) in the mean if there exists a compact set $S \subset \Re^n$ so that for all $x_0 \in S$ there exists a bound μ and a number $\bar{N}(\mu, x_0)$ such that $\|E(x_k) - E(x_e)\| \le \mu$ for all $k > k_0 + \bar{N}$.

Theorem 1: Consider the event sampled NN identifier (19) and the tuning law given by (21). Let the Assumption 3 holds and the NN identifier initial weights $\hat{W}_{I,0} \in \Omega_{W_I}$ with Ω_{W_I} being a compact set. Suppose the vector $\Phi_I(z_k)\bar{u}_k$ satisfies the PE condition. Then, for a positive constant α_I satisfying $0 < \alpha_I < 1/2$ and positive integers N and $\bar{N}$ the identifier NN weight estimation error $\underset{\tau,\gamma}{E}\left(\tilde{W}_{I,k}\right)$ is UB in the mean for all $k_i > k_0 + N$ or, alternatively, $k > k_0 + \bar{N}$ with $\bar{N} > N$.

Proof: Refer to the Appendix.

3.2.2 Stochastic Value Function Approximation: Critic NN Design. In this step we will approximate the value function by using the event sampled critic NN and design the update law to estimate the value function. Recall the event sampled approximation of the value function given in (15). The following assumption holds for the critic NN.

Assumption 4: The target critic NN weight, the activation function and the traditional reconstruction error are upper bounded in the mean such that $\left\|\underset{\tau,\gamma}{E}\{W_V\}\right\| \le W_{V,M}$, $\left\|\underset{\tau,\gamma}{E}\{\phi_V(\bullet)\}\right\| \le \phi_{V,M}$, and $\left\|\underset{\tau,\gamma}{E}\{\varepsilon_V(\bullet)\}\right\| \le \varepsilon_{V,M}$ where $W_{V,M}$, $\phi_{V,M}$, and $\varepsilon_{V,M}$ are positive constants. The gradient of the activation function and traditional reconstruction error are upper bounded in the mean [5], i.e., $\left\|\underset{\tau,\gamma}{E}\{\partial\phi_V(\bullet)/\partial z_{k+1}\}\right\| \le \phi'_{V,M}$ and $\left\|\underset{\tau,\gamma}{E}\{\partial\varepsilon_V(\bullet)/\partial z_{k+1}\}\right\| \le \phi'_{V,M}$. In addition, the activation function is Lipschitz continuous in the compact set for all $z_k \in \Omega_z$. Then, there exists a constant C_{ϕ_V} such that $\left\|\underset{\tau,\gamma}{E}\{\phi_V(z_k)\} - \underset{\tau,\gamma}{E}\{\phi_V(\breve{z}_k)\}\right\| \le C_{\phi_V}\left\|\underset{\tau,\gamma}{E}\{z_k\} - \underset{\tau,\gamma}{E}\{\breve{z}_k\}\right\|$.

The estimated value function by the critic NN in an event sampled context can be written as

$$\hat{V}_k = \underset{\tau,\gamma}{E}\{\hat{W}_{V,k}^T \phi_V(\breve{z}_k)\},\ k_i \le k < k_{i+1}, \tag{24}$$

where $\underset{\tau,\gamma}{E}\{\hat{W}_{V,k}\} \in \Re^{l_V}$ is the estimated critic NN weight matrix. The activation function $\phi_V(\breve{z}_k)$ is selected such that it forms a basis [15] for value function approximation and satisfies $\phi_V(0) = 0$.

Using the estimated value of the value function (24), Equation (6) does not hold any more. Therefore, the temporal difference (TD) or HJB equation error with event sampled estimation of value function is given by

$$\underset{\tau,\gamma}{E}\{e_{V,k}\} = \underset{\tau,\gamma}{E}\{\breve{z}_k^T Q_z \breve{z}_k + u_k^T R_z u_k\} + \underset{\tau,\gamma}{E}\{\hat{W}_{V,k}^T \Delta\phi_V(\breve{z}_k)\}, \tag{25}$$

for $k_i \le k < k_{i+1}$ where $\Delta\phi_V(\breve{z}_k) = \phi_V(\breve{z}_{k+1}) - \phi_V(\breve{z}_k)$.

An augmented HJB error, similar to that in [15], will be used to estimate the critic NN weights without using policy and value iteration. As discussed earlier, the critic NN weights will be updated only at the event sampled instants. Therefore, we define the augmented HJB error at the event sampled instants by defining the new augmented cost-to-go vector Π_k and activation function matrix ψ_k given by $\Pi_k = [r(z_{k_i}, u_{k_i})$ $r(z_{k_{i-1}}, u_{k_{i-1}})$ $\cdots$ $r(z_{k_{i-j-1}}, u_{k_{i-j-1}})] \in \Re^{1\times j}$ and $\psi_k = [\Delta\phi_V(z_{k_i})$ $\Delta\phi_V(z_{k_{i-1}}) \cdots \Delta\phi_V(z_{k_{i-j-1}})] \in \Re^{l_V \times j}$, where $r(\breve{z}_{k_i}, u_{k_i}) = z_{k_i}^T Q_z \breve{z}_{k_i} + u_{k_i}^T R_z u_{k_i}$, $0 < j < i$. Then the augmented HJB error $\Sigma_{V,k}$ is given by

$$\underset{\tau,\gamma}{E}\{\Sigma_{V,k}\} = \underset{\tau,\gamma}{E}\{\Pi_k\} + \underset{\tau,\gamma}{E}\{\hat{W}_{V,k}^T \psi_k\},\ k = k_i. \tag{26}$$

It is clear that convergence of $\underset{\tau,\gamma}{E}\{\Sigma_{V,k}\}$ will ensure convergence of $\underset{\tau,\gamma}{E}\{e_{V,k}\}$. Similar to the identifier, the critic NN weight update law is selected such that it will force the augmented HJB error $\underset{\tau,\gamma}{E}\{\Sigma_{V,k}\}$ close to zero. Define the update law for the critic NN weight as

$$\underset{\tau,\gamma}{E}\{\hat{W}_{V,k}\} = \underset{\tau,\gamma}{E}\left\{\hat{W}_{V,k-1} - \left(\chi_k \alpha_V \psi_{k-1} \Sigma_{V,k-1}^T \Big/ \left\|\psi_{k-1}^T \psi_{k-1} + I\right\|\right)\right\}, \tag{27}$$

where $\alpha_V > 0$ is the critic NN learning gain and χ_k is the event indicator function as defined earlier in (22).

Remark 2: Similar to the NN identifier, the critic NN weights are tuned only at the event sampling instants. Further, the computation of the augmented HJB error, $\Sigma_{V,k-1}$, at $k = k_i$ requires the state vectors z_{k_i} and z_{k_i-1}. Therefore, both the state vectors x_{k_i} and x_{k_i-1} are packetized together and transmitted to the controller at the event sampled instants, $k = k_i$, $\forall i = 1, 2, \cdots$ to form z_{k_i} and z_{k_i-1}.

Finally, using (27) with a forwarded time instant and the definition of critic NN weight estimation error, i.e., $\underset{\tau,\gamma}{E}\{\tilde{W}_{V,k}\} = \underset{\tau,\gamma}{E}\{W_V - \hat{W}_{V,k}\}$, the critic NN weight estimation error dynamics can be represented as

$$\underset{\tau,\gamma}{E}\{\tilde{W}_{V,k+1}\} = \underset{\tau,\gamma}{E}\left\{\tilde{W}_{V,k} + \left(\chi_k \alpha_V \psi_k \Sigma_{V,k}^T \Big/ \left\|\psi_k^T \psi_k + I\right\|\right)\right\}, \tag{28}$$

In terms of the critic NN weight estimation error, $\tilde{W}_{V,k}$, the value function estimation error, $e_{V,k}$, can be rewritten by subtracting (16) from (25)

$$\begin{aligned}\underset{\tau,\gamma}{E}\{e_{V,k}\} &= \underset{\tau,\gamma}{E}\{\breve{z}_k^T Q_z \breve{z}_k - z_k^T Q_z z_k\} - \underset{\tau,\gamma}{E}\{\tilde{W}_{V,k}^T \Delta\phi_V(z_k)\} + \underset{\tau,\gamma}{E}\{\hat{W}_{V,k}^T (\Delta\phi_V(\breve{z}_k) - \Delta\phi_V(z_k)\} \\ &+ \underset{\tau,\gamma}{E}\{\Delta\varepsilon_V(z_k)\},\ k_i \le k < k_{i+1}.\end{aligned} \tag{29}$$

At the event sampled instants with $\breve{z}_k = z_k$, the HJB error (29) can be written as $\underset{\tau,\gamma}{E}\{e_{V,k}\} = -\underset{\tau,\gamma}{E}\{\tilde{W}_{V,k}^T \Delta\phi_V(z_k) + \Delta\varepsilon_V(z_k)\}$, $k = k_i$. Thus, the augmented error at the event sampled instants can be expressed as

$$\underset{\tau,\gamma}{E}\{\Sigma_{V,k}\} = -\underset{\tau,\gamma}{E}\{\tilde{W}_{V,k}^T \psi_k - \Xi_{V,k}\},\ k = k_i\ , \tag{30}$$

where $\Xi_{V,k} = [\Delta\varepsilon_V(z_{k_i})\ \ \Delta\varepsilon_V(z_{k_{i-1}})\ \cdots\ \Delta\varepsilon_V(z_{k_{i-j-i}})]$ is the augmented reconstruction error and satisfies $\left\|\Xi_{V,k}\right\| \le \Xi_{V,M}$.

The next lemma claims the boundedness in the mean of the critic NN weight estimation error.

Lemma 1: Consider the critic NN (24) and its weight tuning law (27). Let the Assumption 4 holds and the critic NN initial weights, $\hat{W}_{V,0} \in \Omega_{W_V}$, with Ω_{W_V} being in a compact set. Suppose the activation function $\phi_V(\breve{z}_k)$ satisfies the PE condition. Then, for a constant learning gain α_V satisfying $0 < \alpha_V < 1/2$, and positive integers N and $\bar{N}$, the critic NN weight estimation error $\underset{\tau,\gamma}{E}\{\tilde{W}_{V,k}\}$ is UB in the mean for all $k_i > k_0 + N$ or, alternatively, $k > k_0 + \bar{N}$ for $\bar{N} > N$.

Proof: Refer to the Appendix.

It only remains to approximate the control input using the actor NN and presented in the next subsection.

3.2.3 Control Input Approximation: Actor NN Design. The optimal control input with event sampled NN approximation can be written as

$$u_k^* = \underset{\tau,\gamma}{E}\{W_u^T \varphi_u(\breve{z}_k) + \varepsilon_{e,u}(\breve{z}_k, e_{ET,k})\},\ k_i \le k < k_{i+1} \tag{31}$$

where $\underset{\tau,\gamma}{E}\{W_u\} \in \Re^{l_u \times m}$ is the unknown constant target NN weights, $\underset{\tau,\gamma}{E}\{\varphi_u(\breve{z}_k)\} \in \Re^{l_u}$ is the activation function and $\varepsilon_{e,\mathrm{u}}(\breve{z}_k, e_{ET,k}) = \underset{\gamma,\tau}{E}\{W_u^T[\varphi_u(\breve{z}_k + e_{ET,k}) - \varphi_u(\breve{z}_k)] + \varepsilon_u(\breve{z}_k + e_{ET,k})\}$ is the event sampled reconstruction error where $\underset{\tau,\gamma}{E}\{\varepsilon_u(z_k)\} \in \Re^m$ is the traditional reconstruction error with l_u as the number of neurons. The following assumption holds for the actor NN.

Assumption 5: The target NN weight, the activation function and the traditional reconstruction error are upper bounded in the mean such that $\left\|\underset{\tau,\gamma}{E}\{W_u\}\right\| \le W_{u,M}$, $\left\|\underset{\tau,\gamma}{E}\{\varphi_u(\bullet)\}\right\| \le \varphi_{u,M}$, and $\left\|\underset{\tau,\gamma}{E}\{\varepsilon_u(\bullet)\}\right\| \le \varepsilon_{u,M}$ where $W_{u,M}$, $\varphi_{u,M}$, and $\varepsilon_{u,M}$ are positive constants. The actor NN activation function $\varphi_u(z_k) \in \Re^{l_u}$ is Lipschitz continuous in a compact set for all $z_k \in \Omega_z$ and satisfies $\left\|\underset{\tau,\gamma}{E}\{\varphi_u(z_k)\} - \underset{\tau,\gamma}{E}\{\varphi_u(\breve{z}_k)\}\right\|$ $\le C_{\varphi_u}\left\|\underset{\tau,\gamma}{E}(z_k) - \underset{\tau,\gamma}{E}(\breve{z}_k)\right\| = C_{\varphi_u}\left\|\underset{\tau,\gamma}{E}\{e_{ET,k}\}\right\|$ where $C_{\varphi_u} > 0$ is a computable constant.

The ideal control input $u_{V,k}^*$ which minimizes the value function (15) can also be computed by computing the gradient of (15) and given by

$$u_{V,k}^* = -\frac{1}{2}\underset{\tau,\gamma}{E}\{R_z^{-1}G^T(z_k)[(\partial\phi_V(\breve{z}_{k+1})/\partial z_{k+1})W_{V,k} + \partial\varepsilon_{e,V}^T(\breve{z}_{k+1}, e_{ET,k+1})/\partial z_{k+1})]\} \tag{32}$$

The control input u_k^*, in (31) is equal to the control input $u_{V,k}^*$ in (32), and, hence, it can be written as

$$\begin{aligned}&\underset{\tau,\gamma}{E}\{W_u^T\varphi_u(\breve{z}_k) + (1/2)R_z^{-1}G^T(z_k)[\partial\phi_V(\breve{z}_{k+1})/\partial z_{k+1})W_{V,k} + \partial\varepsilon_{e,V}^T(\breve{z}_{k+1}, e_{ET,k+1})/\partial z_{k+1}] \\ &+\varepsilon_{e,u}(z_k, e_{ET,k})\} = 0\end{aligned} \tag{33}$$

The estimated control input using the actor NN can be represented as

$$u_k = \underset{\tau,\gamma}{E}\{\hat{W}_{u,k}^T \varphi_u(\breve{z}_k)\},\ k_i \le k < k_{i+1}, \tag{34}$$

where $\underset{\tau,\gamma}{E}(\hat{W}_{u,k}) \in \Re^{l_u \times m}$ is the estimated actor NN weight matrix. The activation function $\varphi_u(\breve{z}_k)$ is selected such that it forms basis for approximation of the control input and satisfies $\varphi_u(0) = 0$.

The estimated control input which minimizes the estimated value function (24) can also be computed by taking the gradient of the estimated value function (24) and given by

$$\begin{aligned} u_{V,k} &= \underset{\tau,\gamma}{E}\left(-(1/2) R_z^{-1} \hat{G}^T(\breve{z}_k)(\partial \hat{V}_{k+1} / \partial z_{k+1})\right) \\ &= \underset{\tau,\gamma}{E}\{-(1/2) R_z^{-1} \hat{G}^T(\breve{z}_k)(\partial \phi_V(\breve{z}_{k+1}) / \partial z_{k+1}) \hat{W}_{V,k}\} \end{aligned}, \tag{35}$$

where $\hat{G}(\breve{z}_k)$ is computed from the identifier. The estimated control inputs (34) and (35) does not satisfy the relation (33). Define the difference between u_k in (34) and $u_{V,k}$ in (35) as control input estimation error. It is given by

$$\underset{\tau,\gamma}{E}\{e_{u,k}\} = \underset{\tau,\gamma}{E}\{\hat{W}_{u,k}^T \varphi_u(\breve{z}_k) + (1/2) R_z^{-1} \hat{G}^T(\breve{z}_k)((\partial \phi_V(\breve{z}_{k+1}) / \partial z_{k+1}) \hat{W}_{V,k})\},\ \ k_i \le k < k_{i+1}. \tag{36}$$

In order to drive the control input estimation error close to zero, similar to the other two NNs, an event sampled update law for the actor NN is selected as

$$\underset{\tau,\gamma}{E}\{\hat{W}_{u,k}\} = \underset{\tau,\gamma}{E}\left\{\hat{W}_{u,k-1} - \frac{\chi_k \alpha_u \varphi_u(\breve{z}_{k-1}) e_{u,k-1}^T}{\varphi_u^T(\breve{z}_{k-1}) \varphi_u(\breve{z}_{k-1}) + 1}\right\}, \tag{37}$$

where $\alpha_u > 0$ is the actor NN learning gain.

Defining the actor NN weight estimation error as $\underset{\tau,\gamma}{E}\{\tilde{W}_{u,k}\} = \underset{\tau,\gamma}{E}\{W_u - \hat{W}_{u,k}\}$, the actor NN weight estimation error dynamics by using (37) can be written as

$$\underset{\tau,\gamma}{E}\{\tilde{W}_{u,k+1}\} = \underset{\tau,\gamma}{E}\left\{\tilde{W}_{u,k} + \frac{\chi_k \alpha_u \varphi_u(\breve{z}_k) e_{u,k}^T}{\varphi_u^T(\breve{z}_k)\varphi_u(\breve{z}_k)+1}\right\}, \tag{38}$$

Further, subtracting (33) from (36), the control input estimation error can be written as

$$\begin{aligned}\underset{\tau,\gamma}{E}\{e_{u,k}\} = \underset{\tau,\gamma}{E}\{&-\tilde{W}_u^T\varphi_u(z_k) + \hat{W}_{u,k}^T\left(\varphi_u(\breve{z}_k) - \varphi_u(z_k)\right) - (1/2)R_z^{-1}\tilde{G}^T(z_k)(\partial\phi_V(z_{k+1})/\partial z_{k+1})W_V \\ &-(1/2)R_z^{-1}\hat{G}^T(z_k)(\partial\phi_V(z_{k+1})/\partial z_{k+1})\tilde{W}_{V,k} + (1/2)R_z^{-1}\left(\hat{G}^T(\breve{z}_k)(\partial\phi_V(\breve{z}_{k+1})/\partial z_{k+1})\right. \\ &\left.-\hat{G}^T(z_k)(\partial\phi_V(z_{k+1})/\partial z_{k+1})\right)\hat{W}_{V,k} - \varepsilon_{e,k}\}\end{aligned} \tag{39}$$

where $\varepsilon_{e,k} = \varepsilon_u(z_k) + (1/2)R_z^{-1}G^T(z_k)(\partial\varepsilon_u^T(z_k)/\partial z_{k+1})$ satisfying $\|\varepsilon_{e,k}\| \le \varepsilon_{e,M}$ and $\tilde{G}(z_k) = \tilde{G}(z_k) - \hat{G}(z_k)$. We will use (39) for closed-loop stability presented next.

3.3 DESIGN OF EVENT SAMPLING CONDITION AND STABILITY

In this section, the UB of the event sampled closed-loop system is shown by using Lyapunov stability technique and designing an adaptive event sampling condition.

The closed-loop system dynamics by using (4) and control input (34) becomes

$$\begin{aligned}z_{k+1} = {} & F(z_k) + G(z_k)u_k^* - G(z_k)(\underset{\tau,\gamma}{E}\{\tilde{W}_{u,k}^T\varphi_u(z_k) + \underset{\tau,\gamma}{E}\{\hat{W}_{u,k}^T\left(\varphi_u(z_k) - \varphi_u(\breve{z}_k)\right)\} \\ & + \underset{\tau,\gamma}{E}\{\varepsilon_u(z_k)\},\ k_i \le k < k_{i+1}.\end{aligned} \tag{40}$$

Before claiming the main results for the closed-loop event sampled system, the following lemma is necessary.

Lemma 2[4]: Consider the controllable augmented NCS system (4) and the optimal control input u_k^* in (7). Then the closed-loop system satisfies

$$\left\|\underset{\tau,\gamma}{E}\{F(z_k) + G(z_k)u_k^*\}\right\|^2 \le \mu\left\|\underset{\tau,\gamma}{E}\{z_k\}\right\|^2, \tag{41}$$

where $0 < \mu < 1$ is a constant.

Now, we will introduce the adaptive event sampling condition to decide the transmission instants. The system state vector and updated control input are transmitted when the following adaptive event sampling condition is violated. It is given by

$$\sigma_{ET,k}\ D\left(\left\|\underset{\tau,\gamma}{E}\{e_{ET,k}\}\right\|^2\right) \le \rho\left\|\underset{\tau,\gamma}{E}\{z_k\}\right\|^2, \tag{42}$$

where $\sigma_{ET,k} = 12G_M^2 C_{\varphi_u}^2 \left\|\underset{\tau,\gamma}{E}\{\hat{W}_{u,k}\}\right\|^2 + C_{\Phi_I}\left\|\underset{\tau,\gamma}{E}\{\hat{W}_{I,k}\}\right\|^2$ and $\rho = 2\Gamma(1-2\mu)$ with $0<\Gamma<1$, $0<\mu<1/2$, and G_M is the upper bound of the matrix function $G(z_k)$. The operator $D(\bullet)$ is the dead zone operator and defined as

$$D(\bullet) = \begin{cases} \bullet, & \text{if } \left\|\underset{\tau,\gamma}{E}\{z_k\}\right\| > B_z \\ 0, & \text{otherwise} \end{cases} \tag{43}$$

where B_z is the UB for the system state. To ensure the estimated NN weights $\left\|\underset{\tau,\gamma}{E}\{\hat{W}_{u,k}\}\right\|$ and $\left\|\underset{\tau,\gamma}{E}\{\hat{W}_{I,k}\}\right\|$ in (42) are non-zero while evaluating the trigger condition, the previous nonzero values are used for evaluating the event sampling condition when the current estimated values become zero.

Remark 3: As proposed, the event sampling condition (42) is a function of the actor and identifier NN weight estimates. With the update of the NN weights the condition also gets updated. This adaptive condition generates the required number of event sampled instants based on the NN weight estimation error during the initial learning phase. Once the NN weights converge to the UB, the coefficient $\sigma_{ET,k}$ becomes a constant and the event sampling condition becomes same as the traditional event-triggering condition [16]-[18]. In addition, the dead zone operator in the event sampling condition prevents

the unnecessary event sampled instants generated due to the NN reconstruction error once the system state is inside the UB.

Next, the UB in the mean of the closed-loop event-triggered system is claimed in the following theorem.

Theorem 2: Consider the NCS dynamics (4) along with the NN identifier, critic NN and actor NN defined as in (19), (24) and (34) with event sampled weight update laws (21), (27) and (37) respectively. Let the Assumptions 1 through 5 hold and the NN initial weights $\hat{W}_{I,0} \in \Omega_{\hat{W}_I}$, $\hat{W}_{V,0} \in \Omega_{W_V}$ and $\hat{W}_{u,0} \in \Omega_{W_u}$. Suppose the last transmitted state, $\breve{z}_k$, the NN weights $\hat{W}_{I,k}$, $\hat{W}_{V,k}$, $\hat{W}_{u,k}$, and control input, u_k, are updated at the violation of the event sampling condition (42). Then, for learning gains satisfying $0 < \alpha_I < 1/2$, $0 < \alpha_V < 1/2$ and $0 < \alpha_u < 1/4$, and positive integers N and $\bar{N}$ the event sampled closed-loop system state $\underset{\tau,\gamma}{E}\{z_k\}$, identifier estimation error $\underset{\tau,\gamma}{E}\{\tilde{z}_k\}$, identifier, critic NN, and the actor NN weight estimation errors $\underset{\tau,\gamma}{E}\{\tilde{W}_{V,k}\}$, $\underset{\tau,\gamma}{E}\{\tilde{W}_{I,k}\}$, and $\underset{\tau,\gamma}{E}\{\tilde{W}_{u,k}\}$, respectively, are UB in the mean for all sampling instants $k_i > k_0 + N$ or alternatively, $k > k_0 + \bar{N}$ for $\bar{N} > N$. Further, the estimated value function and control policy converge close to their respective optimal values, i.e., $\left\|\hat{V}_k - V_k^*\right\| \leq B_V$ and $\left\|u_k - u_k^*\right\| \leq B_u$ where B_V and B_u are small positive constants.

Proof: Refer to Appendix.

Remark 4: The bound B_z is a function of design parameters and NN reconstruction error. Hence, the bound for the system state can be made small by increasing the number of neurons in the NNs and selecting the learning gains accordingly.

4. SIMULATION RESULTS

In this section, the analytical results are illustrated with a numerical example of second order system. The continuous-time system dynamics chosen for simulation are given by [17]

$$\begin{aligned} \dot{x}_1 &= x_2, \\ \dot{x}_2 &= -x_1 - x_1^3 + u \end{aligned}. \tag{44}$$

The simulation parameters were selected as follows. Sampling time for the sensor sampling is chosen ad $T_s = 0.01\text{sec}$, time varying delay bound chosen as $\bar{d} = 2$, the mean value of the delay is chosen as $\underset{\tau,\gamma}{E}(\tau) = 12ms$. The packet losses follows a Bernoulli distribution with $p = 0.4$. The penalty matrices were selected as $P_z = I_{4\times 4}$ and $R_z = 1$. The critic NN activation function was selected as $\phi_V(\bullet) = \tanh\left(z_{1,k}^2; z_{1,k}z_{2,k}; \cdots; z_{2,k}^2; z_{2,k}z_{3,k}; \cdots; z_{4,k}^2; z_{1,k}^4; z_{1,k}^3 z_{2,k}; \cdots; z_{1,k}z_{2,k}z_{3,k}z_{4,k}; z_{4,k}^4\right) \in \Re^{39}$, the actor NN and identifier NN activation functions respectively are $\Phi_I(\bullet) = \tanh(\bullet)$ and $\varphi_u(\bullet) = \tanh(\bullet)$. Number of neurons for the critic is 39, actor is 15 and identifier is 50. The learning gains were $\alpha_V = 0.04$, $\alpha_u = 0.05$, and $\alpha_I = 0.01$. The parameters for the event-trigger condition were selected as $\mu = 0.45$, $\Gamma = 0.99$, $G_M = 1$ and the Lipschitz constant for identifier and the actor NN activation function is computed to be 5.2 and 16.9 respectively. All the three NN weights were initialized at random from a uniform distribution in the interval $\begin{bmatrix} 0 & 1 \end{bmatrix}$. Further, the initial state vector is taken as $x_0 = \begin{bmatrix} 2 & -1 \end{bmatrix}^T$ and for the identifier as $\hat{z}_0 = \begin{bmatrix} 1 & -3 & 0 & 0 \end{bmatrix}^T$. The Monte Carlo simulation is run using the same initial condition

with randomly generated delays and packet losses. The simulation results are given in Figures 2 through 5.

Figure 2 (a) and (b) shows the regulation of the systems state vector and convergence of the control policy close to zero. The stochastic optimal controller is able to regulate the system in the presence of the networked induced time-varying delays and packet losses.

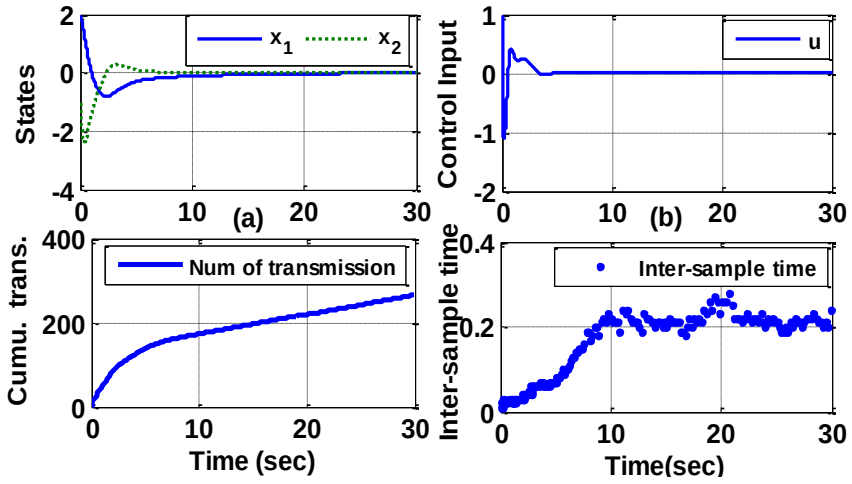

Figure 2. Time history of (a) the system state; (b) control input; (c) cumulative number of transmission or event sampled instants; (d) inter-sample times.

With respect to the number of transmissions, it is clear from Figure 2 (c) that a total of 285 transmissions occurred out of 3000 sampling instants. This shows the reduction in bandwidth and also computation since the controller is executed at the event sampled instants only. It is important to note that the number of event sampled instants varies with initial values of the NNs. Further, the aperiodic inter-sample times are depicted in Fig. 2 (d). Once the NN weights converge close to the target values the inter-sample times or the transmission intervals are elongated.

The event sampling condition (42) is plotted in Fig. 3(a) shows the evolution of the event sampling error (zoomed figure) between event sampled instants. It is observed that there is a frequent transmission during the initial learning phase due to large initial approximation errors. The convergence of the HJB error close to zero is shown in Fig. 3(b). This indicates the near optimality of the control input.

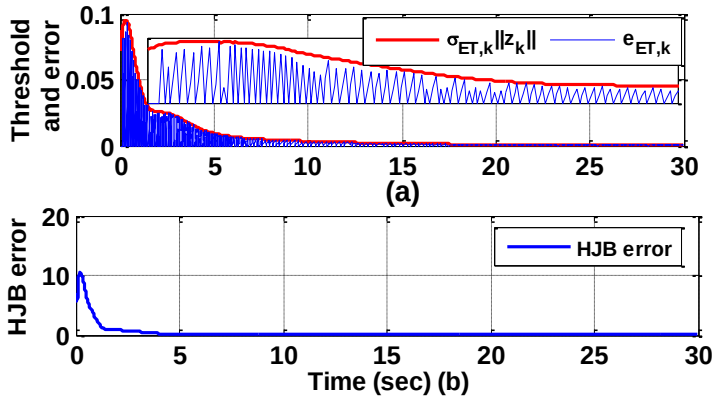

Figure 3. Evolution of (a) the event sampling threshold and event sampling error and (b) HJB error.

Considering a data packet of 8 bit data, the comparison of the data rate for a periodic system and event sampled system is shown in Fig. 4. An average data rate of 270 bits/ sec for event sampled system is observed where as in periodic 800 bits/ sec. It shows a saving of approximately 66% of the bandwidth usage. The reduction in computation in terms of the addition and multiplication for the proposed design when compared to traditional periodic implementation is shown in Table 1. A reduction of 27% was observed.

Further, the convergence of the three NN weights is shown in Fig. 5. The norm of the NN weight estimates become constant implying the convergence of the estimates.

This further implies the boundedness of the NN weight estimation errors and their convergence to the ultimate bound.

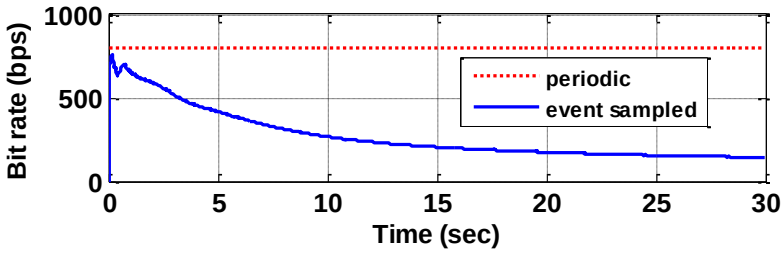

Figure 4. Comparison of data rate (bits/sec) between periodic and event sampled system.

Table 1. Comparison of computational load between traditional periodic sampled and event sampled system.

System			Traditional periodic sampled	Event-based non-periodic sampled
Samping instants			3000	285
Number of additions and Multiplications at every sampling instant	Critic and update law		8	8
	Actor NN and update law		8	8
	Identifier and update		12	12
	Trig. Mechnaism	Critic and update law	0	8
		Actor NN and update law	0	8
		Identifier and update	0	12
		Event sampling condition (periodic)	0	15
Total number of Computation			84000	60960

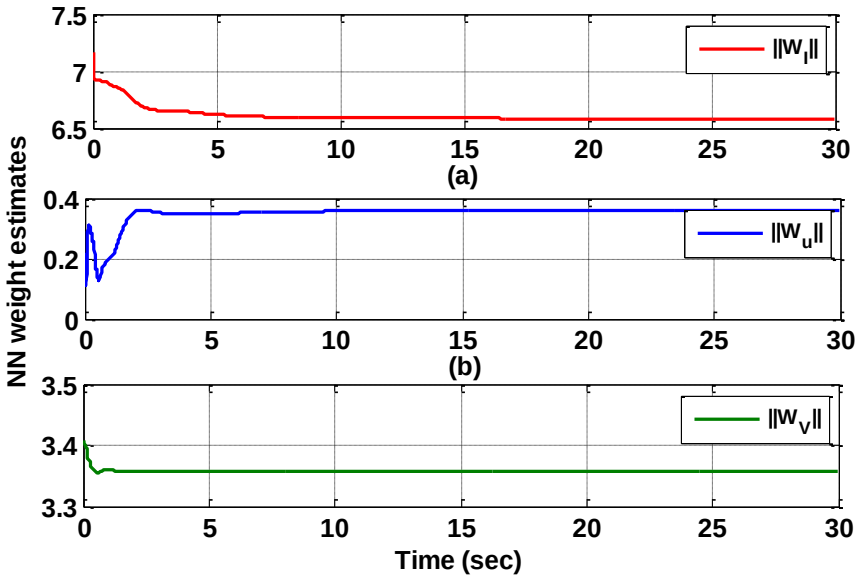

Figure 5. Convergence of the NN weight estimates.

5. CONCLUSIONS

In this paper, a stochastic event sampled ADP based near optimal regulator design is presented in the presence of network induced delays and packet losses. The additional error introduced in the system due to event sampled transmission was taken care of by designing the adaptive event sampling condition. The adaptive event sampling condition is found to provide reasonable accuracy in approximation while ensuring stability. A reduction is transmission is observed along with near optimal performance. Initial transmissions found to be higher due to large NNs weight estimation error. Further, it was observed that different initial condition and learning gain resulted in different number of event sampled instants and transmission to achieve the approximation accuracy. Finally, the simulation results validated the analytical design. Furthermore, the cost function considered in this case only penalizes the system state and control input. It will be interesting to optimize the transmission instants and included in our future research.

6. REFERENCES

[1] G. C. Walsh, Ye Hong, and L. G. Bushnell, "Stability analysis of networked control systems," *IEEE Transactions on Control Systems Technology,* vol.10, no. 3, pp. 438-446, May 2002.

[2] L. W. Liou and A. Ray, "A stochastic regulator for integrated communication and control systems: part I—formulation of control law," *ASME Journal of Dynamic Systems, Measurement, and Control*, vol. 113, no. 4, pp. 604–611, Jan. 1991.

[3] Hu Shousong and Zhu Qixin, "Stochastic optimal control and analysis of stability of networked control systems with long delay," *Automatica,* vol. 39, no. 11, pp. 1877-1884, Nov. 2003.

[4] H. Xu, S. Jagannathan, and F. L. Lewis, "Stochastic optimal control of unknown linear networked control system in presence of random delays and packet losses," *Automatica,* vol. 48, pp. 1017–1030, Jun. 2012.

[5] Hao Xu and S. Jagannathan, " Stochastic optimal controller design for uncertain nonlinear networked control system via neuro dynamic programming," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 24, no. 3, pp. 471 – 484, Jan. 2013.

[6] H. Xu and S. Jagannathan, "Neural network based finite horizon stochastic optimal controller design for nonlinear networked control systems," in *Proceedings of the IEEE International Joint Conference on Neural Network*, Beijing, China, Jul. 2013.

[7] M. C. F. Donkers, *Networked and Event-Triggered Control Systems.* PhD dissertation, Eindhoven University of Technology, Netherlands, 2011.

[8] D. P. Bertsekas, *Dynamic Programming and Optimal Control,* 3nd ed. Belmonth, MA: Athena Scientific, 2000.

[9] D. V. Prokhorov and D. Wunsch, "Adaptive Critic Designs," *IEEE Transactions on Neural Networks*, vol 8, no. 5 , pp. 997–1007, Sep. 1997.

[10] H. Zhang, Y. Luo, and D. Liu, "Neural-network-based near-optimal control for a class of discrete-time affine nonlinear systems with control constraints," *IEEE Transactions on Neural Network*, vol. 20, no. 9, pp. 1490–1503, Sep. 2009.

[11] D. Liu and Q. Wei, "Finite-approximation-error based optimal control approach for discrete-time nonlinear systems," *IEEE Transactionson Cybernetics*, vol. 43, no. 2, pp. 779-789, Apr. 2013.

[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 Transactions on Systtem, Man, and Cybernetiucs- B,* vol. 38, no. 4, pp. 937–942, Aug. 2008.

[13] A. Al-Tamimi and F. L. Lewis, "Discrete-time nonlinear HJB solution using approximate dynamic programming: Convergence proof," *IEEE Transactions on System, Man., and Cybernetics- B,* vol. 38, no. 4, pp. 943–949, Aug. 2008.

[14] C. Zheng and S. Jagannathan, "Generalized Hamilton–Jacobi–Bellman formulation-based neural network control of affine nonlinear discrete time systems," *IEEE Transactions on Neural Networks,* vol. 19, no. 1, pp. 90–106,Jan. 2008.

[15] 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 Transactions on Neural Networks and Learning Systems,* vol. 23, no. 7, pp. 1118-1129, May. 2012.

[16] X. Wang and M. Lemmon, "Event triggering in distributed networked control systems," *IEEE Transactions on Automatic Control,* vol. 56, no. 3, pp. 586–601, Mar. 2011.

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

[18] 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.

[19] R. Cogill, "Event-based control using quadratic approximate value functions," in *Proceedings of the 48th Joint IEEE Conference on Decision and Control and Chinese Control Conference*, Shanghai, China, Dec. 2009, pp. 5883–5888.

[20] A. Molin and S. Hirche, "On the optimality of certainty equivalence for event-triggered control systems," *IEEE Transactions on Automatic Control,* vol. 58, no. 2, pp. 470-474, Feb. 2013.

[21] M. Rabi, K. H. Johansson, and M. Johansson, "Optimal stopping for event-triggered sensing and actuation," In *Proceedings of the 47th IEEE Conference on Decision and Control,* Cancun, Mexico, Dec. 2008, pp.3607-3612.

[22] O.C. Imer and T. Basar, "To measure or to control: optimal control with scheduled measurements and controls," in *Proceedings* of *the American Control Conference,* Minneapolis, MN, USA, Jun. 2006, pp. 14-16.

[23] A. Sahoo, Hao Xu, and S. Jagannathan, "Event-based optimal regulator design for nonlinear networked control systems," in *proceedings of the IEEE symposium on Adaptive Dynamic Programming and Reinforcement Learning*, Orlando, FL, Dec. 2014, pp. 1-8.

[24] S. Jagannathan, *Neural Network Control of Nonlinear Discrete-time System.* Boca Raton, FL: CRC Press, 2006.

APPENDIX

Proof of Lemma 1: The proof is carried out considering two cases of the event sampling scenario, i.e., event sampled instants (identifier NN weights are updated) and inter-sample times (identifier NN weights are not updated). We will evaluate a common Lyapunov function for both the cases and finally combine them to show UB of the identifier NN weight estimation error.

Case I: At event sampling instants ($k = k_i,\ \forall 1 = 1, 2, \cdots$)

At the event sampling instants $\bar{z}_k = z_k$, $\chi_k = 1$, $k = k_i$, $\forall 1 = 1, 2, \cdots$. Therefore, the NN identifier weight estimation error (23) at $k = k_i$ becomes

$$\underset{\tau,\gamma}{E}\{\tilde{W}_{I,k+1}\} = \underset{\tau,\gamma}{E}\left\{\tilde{W}_{I,k} - \frac{\alpha_I \Phi_I(z_k)\bar{u}_k \tilde{z}_{I,k+1}^T}{(\Phi_I(z_k)\bar{u}_k)^T(\Phi_I(z_k)\bar{u}_k)+1}\right\}, k = k_i, \tag{A.1}$$

where the identification error dynamics (20) at event sampled instants is given by

$$\tilde{z}_{k+1} = \tilde{W}_{I,k}^T \Phi_I(z_k)\bar{u}_k + \bar{\varepsilon}_{I,k},\ k = k_i \tag{A.2}$$

Consider the Lyapunov function candidate given by

$$L_{I,k} = tr\left\{\underset{\tau,\gamma}{E}\{\tilde{W}_{I,k}^T \tilde{W}_{I,k}\}\right\}. \tag{A.3}$$

The first difference of (A.3) along the dynamics (A.1) can be represented as

$$\begin{aligned}
\Delta L_{I,k} &= tr\left(\underset{\tau,\gamma}{E}\{\tilde{W}_{I,k+1}^T \tilde{W}_{I,k+1}\}\right) - tr\left(\underset{\tau,\gamma}{E}\{\tilde{W}_{I,k}^T \tilde{W}_{I,k}\}\right) \\
&= -tr\left(\underset{\tau,\gamma}{E}\left\{\frac{2\alpha_I \tilde{W}_{I,k}^T \Phi_I(z_k)\bar{u}_k \tilde{z}_{I,k+1}^T}{(\Phi_I(z_k)\bar{u}_k)^T(\Phi_I(z_k)\bar{u}_k)+1}\right\}\right) \\
&+ tr\left(\underset{\tau,\gamma}{E}\left\{\frac{\alpha_I^2 \tilde{z}_{I,k+1}(\Phi_I(z_k)\bar{u}_k)^T(\Phi_I(z_k)\bar{u}_k)\tilde{z}_{I,k+1}^T}{((\Phi_I(z_k)\bar{u}_k)^T(\Phi_I(z_k)\bar{u}_k)+1)^2}\right\}\right), k = k_i,\ \forall 1 = 1, 2, \cdots.
\end{aligned}$$

Substituting the identification error dynamics (A.2) and applying Cauchy-Schwartz (C-S) inequality, $(a+b)^T(a+b) \le 2a^T a + 2b^T b$, the first difference is bounded above as

$$\Delta L_{I,k} \le -\bar{\Phi}_{I,m}^2 \alpha_I (1-2\alpha_I) \left\| \underset{\tau,\gamma}{E}\{\tilde{W}_{I,k}\} \right\|^2 + (1+2\alpha_I)\alpha_I \bar{\varepsilon}_{I,M}^2 . \tag{A.4}$$

for $k = k_i$ where $0 < \bar{\Phi}_{I,m}^2 \le \min_k \left\{ \frac{\left\| (\Phi_I(z_k)\bar{u}_k)(\Phi_I(z_k)\bar{u}_k)^T \right\|}{(\Phi_I(z_k)\bar{u}_k)^T(\Phi_I(z_k)\bar{u}_k)+1} \right\}$ is satisfied due to PE condition as discussed in Remark 1, $0 < \alpha_I < 1/2$ and $\left\| \bar{\varepsilon}_{I,k} \right\| \le \bar{\varepsilon}_{I,M}$ for identification. From (A.4), the first difference $\Delta L_{I,k} < 0$ is negative as long as

$$\left\| \underset{\tau,\gamma}{E}\{\tilde{W}_{I,k}\} \right\| > \sqrt{(1+2\alpha_I)\bar{\varepsilon}_{I,M}^2 / \bar{\Phi}_{I,m}^2 (1-2\alpha_I)} = B_{ub}^{\tilde{W}_I} .$$

By Lyapunov theorem [24] the NN identifier weight estimation error, $\underset{\tau,\gamma}{E}\{\tilde{W}_{I,k}\}$, is ultimately bounded (UB) in the mean at the event sampled instnats.

Case II: During inter-sample times ($k_i < k < k_{i+1}$)

In this case, the identifier weights are not tuned and held at the previous values. This implies $\chi_k = 0$, $k_i < k < k_{i+1}$. Therefore, the identifier NN weight estimation error dynamics can be rewritten as

$$\underset{\tau,\gamma}{E}\{\tilde{W}_{I,k+1}\} = \underset{\tau,\gamma}{E}\{\tilde{W}_{I,k}\}, k_i < k < k_{i+1} . \tag{A.5}$$

Consider the same Lyapunov function (A.3) in Case I. The first difference using (A.5) can be written as

$$\Delta L_{I,k} = tr\left(\underset{\tau,\gamma}{E}\{\tilde{W}_{I,k+1}^T \tilde{W}_{I,k+1}\} \right) - tr\left(\underset{\tau,\gamma}{E}\{\tilde{W}_{I,k}^T \tilde{W}_{I,k}\} \right) = 0 . \tag{A.6}$$

From (A.6) , the first difference, $\Delta L_{I,k} = 0$ This implies the NN $\underset{\tau,\gamma}{E}\{\tilde{W}_{I,k}\}$ is held for $k_i < k < k_{i+1}, \forall 1 = 1, 2, \cdots$.

Now combining both Cases, we will show that $\underset{\tau,\gamma}{E}\{\tilde{W}_{I,k}\}$ is UB in the mean. Since, the initial NN weights $\hat{W}_{I,0} \in \Omega_{W_I}$ are bounded and the target NN weights W_I are

also bounded from Assumption 3, $\tilde{W}_{I,0}$ is bounded. Further, from Case I, $\underset{\tau,\gamma}{E}\{\tilde{W}_{I,k_i+1}\} \le \underset{\tau,\gamma}{E}\{\tilde{W}_{I,k_i}\}$ and from Case II, $\underset{\tau,\gamma}{E}\{\tilde{W}_{I,k_i+1}\} = \underset{\tau,\gamma}{E}\{\tilde{W}_{I,k}\} = \underset{\tau,\gamma}{E}\{\tilde{W}_{I,k_{i+1}}\}$, $k_i < k < k_{i+1}$. From both the cases, we have $\underset{\tau,\gamma}{E}\{\tilde{W}_{I,k_{i+1}}\} \le \underset{\tau,\gamma}{E}\{\tilde{W}_{I,k_i}\}$. This implies $\underset{\tau,\gamma}{E}\{\tilde{W}_{I,k_i}\} \to B_{ub}^{\tilde{W}_I}$ is for all $k_i > k_0 + N$ where N being a positive integer. Since k_i is a subsequence of $k \in \mathbb{N}$, there exist a positive integer satisfying $\bar{N} > N$ such that $\underset{\tau,\gamma}{E}\{\tilde{W}_{I,k}\} \to B_{ub}^{\tilde{W}_I}$ for all time $k > k_0 + \bar{N}$. Consequently, $\underset{\tau,\gamma}{E}\{\tilde{W}_{I,k}\}$ is UB in the mean with an ultimate bound $B_{ub}^{\tilde{W}_I}$. ■

Proof of Lemma 2: The proof for UB in the mean is shown similar to the NN identifier by considering both event sampled instants and the inter-sample times with a common Lyapunov function.

Case 1: At event sampled instants ($k = k_i$, $\forall 1 = 1, 2, \cdots$)

At the event sampled instants, $\chi_k = 1$, $\breve{z}_k = z_k$, $k = k_i$, $\forall 1 = 1, 2, \cdots$. Therefore, the critic NN weight estimation error dynamics (28) becomes

$$\underset{\tau,\gamma}{E}\{\tilde{W}_{V,k}\} = \underset{\tau,\gamma}{E}\left\{\tilde{W}_{V,k} + \left(\alpha_V \psi_k \Sigma_{V,k}^T \Big/ \left\|\psi_k^T \psi_k + I\right\|\right)\right\},\ k = k_i. \tag{A.7}$$

where $\Sigma_{V,k}$ is given in (30).

Consider the Lyapunov function candidate given by

$$L_{V,k} = \underset{\tau,\gamma}{E}\{\tilde{W}_{V,k}^T \tilde{W}_{V,k}\}. \tag{A.8}$$

The first difference $\Delta L_{V,k}$ along with the critic NN weight estimation error dynamics (A.8) can be written as

$$\Delta L_{V,k} = \underset{\tau,\gamma}{E}\{\tilde{W}_{V,k+1}^T \tilde{W}_{V,k+1}\} - \underset{\tau,\gamma}{E}\{\tilde{W}_{V,k}^T \tilde{W}_{V,k}\}. \tag{A.9}$$

$$= 2\underset{\tau,\gamma}{E}\left\{\frac{\alpha_V \tilde{W}_{V,k}^T \psi_k \Sigma_{V,k}^T}{\left\|\psi_k^T\psi_k + I\right\|}\right\} + \underset{\tau,\gamma}{E}\left\{\frac{\alpha_V^2 \Sigma_{V,k}\psi_k^T\psi_k\Sigma_{V,k}^T}{\left\|\psi_k^T\psi_k + I\right\|^2}\right\}$$

Substituting the augmented HJB error (30), the first difference leads to

$$\Delta L_{V,k} = -2\underset{\tau,\gamma}{E}\left\{\left(\alpha_V \tilde{W}_{V,k}^T\psi_k\psi_k^T\tilde{W}_{V,k} - \alpha_V\tilde{W}_{V,k}^T\psi_k\Xi_k^T\right)\Big/\left\|\psi_k^T\psi_k + I\right\|\right\}$$
$$+\underset{\tau,\gamma}{E}\left\{\left(\alpha_V^2(\tilde{W}_{V,k}^T\psi_k - \Xi_{V,k})\psi_k^T\psi_k(\tilde{W}_{V,k}^T\psi_k - \Xi_{V,k})^T\right)\Big/\left\|\psi_k^T\psi_k + I\right\|^2\right\}$$

Observe that $0 < \bar{\psi}_{\min} \le \left\|\psi_k\psi_k^T\right\|/\left\|\psi_k^T\psi_k + I\right\| < 1$ satisfied due to PE condition. Applying C-S inequality with simple mathematical operations, the first difference is upper bounded by

$$\Delta L_{V,k} \le -(1-2\alpha_V)\alpha_V\underset{\tau,\gamma}{E}\left\{\tilde{W}_{V,k}^T\psi_k\psi_k^T\tilde{W}_{V,k}\Big/\left\|\psi_k^T\psi_k + I\right\|\right\} + 2\alpha_V^2\underset{\tau,\gamma}{E}\left\{\Xi_{V,k}\Xi_{V,k}^T\right\} + \alpha_V\underset{\tau,\gamma}{E}\left\{\Xi_{V,k}\Xi_{V,k}^T\right\}$$

Using Frobenius norm and collecting the similar terms, the first difference is upper bounded as

$$\Delta L_{V,k} \le -\alpha_V(1-2\alpha_V)\bar{\psi}_{\min}\left\|\underset{\tau,\gamma}{E}\{\tilde{W}_{V,k}\}\right\|^2 + (1+2\alpha_V)\alpha_V\Xi_{V,M}^2 . \tag{A.10}$$

From (A.10), the first difference $\Delta L_{V,k} < 0$ as long as

$$\left\|\underset{\tau,\gamma}{E}\{\tilde{W}_{V,k}\}\right\| > \sqrt{(1+2\alpha_V)\Xi_{V,M}^2/(1-2\alpha_V)} = B_{ub}^{\tilde{W}_V} .$$

By Lyapunov theorem [24] the critic NN weight estimation error, $\underset{\tau,\gamma}{E}\{\tilde{W}_{V,k}\}$, is ultimately bounded (UB) at the event sampled instants.

Case II: During inter-sample times ($k_i < k < k_{i+1}$)

In this case, the critic NN weights are not updated since the indicator function $\chi_k = 0$, $k_i < k < k_{i+1}$. Considering the same Lyapunov function (A.8) as in Case 1, the first difference

$$\Delta L_{V,k} = \underset{\tau,\gamma}{E}\{\tilde{W}_{V,k+1}^T \tilde{W}_{V,k+1}\} - \underset{\tau,\gamma}{E}\{\tilde{W}_{V,k}^T \tilde{W}_{V,k}\} = 0. \quad \text{(A.11)}$$

From (A.11), the critic NN weight estimation error $\underset{\tau,\gamma}{E}\{\tilde{W}_{V,k}\}$ remains constant during the inter-sample times.

Now we will combine both Cases to show the UB in the mean. From Case I, we have $\underset{\tau,\gamma}{E}\{\tilde{W}_{V,k_i+1}\} \leq \underset{\tau,\gamma}{E}\{\tilde{W}_{V,k_i}\}$ and from Case II $\underset{\tau,\gamma}{E}\{\tilde{W}_{V,k_i+1}\} = \underset{\tau,\gamma}{E}\{\tilde{W}_{V,k}\} = \underset{\tau,\gamma}{E}\{\tilde{W}_{V,k_{i+1}}\}$, $k_i < k < k_{i+1}$. From both the cases, $\underset{\tau,\gamma}{E}\{\tilde{W}_{V,k_{i+1}}\} \leq \underset{\tau,\gamma}{E}\{\tilde{W}_{V,k_i}\}$, $\forall i = 1,2,\cdots$. Therefore, $\underset{\tau,\gamma}{E}\{\tilde{W}_{V,k_i}\} \to B_{ub}^{\tilde{W}_V}$ for all $k_i > k_0 + N$ where N is a positive integer. Since k_i is a subsequence of $k \in \mathbb{N}$, by extension $\underset{\tau,\gamma}{E}\{\tilde{W}_{V,k}\} \to B_{ub}^{\tilde{W}_V}$ for all time $k > k_0 + \bar{N}$ where $\bar{N} > N$ is a positive integer. Consequently, $\underset{\tau,\gamma}{E}\{\tilde{W}_{V,k}\}$ is UB in the mean with an ultimate bound $B_{ub}^{\tilde{W}_V}$. ■

Proof of Theorem 1: The proof of UB in the mean for the event sampled closed-loop system is carried out by evaluating a common Lyapunov function both at event sampled instants and inter-sample times.

Case I: At event sampled instants ($k = k_i$, $\forall 1,2,\cdots$)

At the event sampled instants we have $\bar{z}_k = z_k$, $e_{ET,k} = 0$ and $\chi_k = 1$, $k = k_i$, $\forall 1 = 1,2,\cdots$. By using these facts, the control input error $\underset{\tau,\gamma}{E}\{e_{u,k}\}$ from (39) becomes

$$\underset{\tau,\gamma}{E}\{e_{u,k}\} = \underset{\tau,\gamma}{E}\left\{-\tilde{W}_u^T \varphi_u(z_k) - \frac{1}{2} R_z^{-1} \tilde{G}^T(z_k) \frac{\partial \phi_V(z_{k+1})}{\partial z_{k+1}} W_V - \frac{1}{2} R_z^{-1} \hat{G}^T(z_k) \frac{\partial \phi_V(z_{k+1})}{\partial z_{k+1}} \tilde{W}_{V,k} - \varepsilon_{e,k}\right\}. \quad \text{(A.12)}$$

Then, the actor NN weight estimation error dynamics at the event sampling instants, by using (A.12), can be written as

$$\underset{\tau,\gamma}{E}\{\tilde{W}_{u,k+1}\}=\underset{\tau,\gamma}{E}\{\tilde{W}_{u,k}\}-\underset{\tau,\gamma}{E}\left\{\frac{\alpha_u\varphi_u(z_k)(\tilde{W}_u^T\varphi_u(z_k))^T}{\varphi_u^T(z_k)\varphi_u(z_k)+1}\right\}$$
$$-\frac{1}{2}\underset{\tau,\gamma}{E}\left\{\frac{\alpha_u\varphi_u(z_k)}{\varphi_u^T(z_k)\varphi_u(z_k)+1}\left(R_z^{-1}\tilde{G}^T(z_k)\frac{\partial\phi_V(z_{k+1})}{\partial z_{k+1}}W_V\right)^T\right\} \tag{A.13}$$
$$-\frac{1}{2}\underset{\tau,\gamma}{E}\left\{\frac{\alpha_u\varphi_u(z_k)}{\varphi_u^T(z_k)\varphi_u(z_k)+1}\left(R_z^{-1}\hat{G}^T(z_k)\frac{\partial\phi_V(z_{k+1})}{\partial z_{k+1}}\tilde{W}_{V,k}\right)^T\right\}-\underset{\tau,\gamma}{E}\{\varepsilon_{e,k}\}.$$

Further, the closed-loop dynamics from (40) at the event sampling instants can be written as

$$z_{k+1}=F(z_k)+G(z_k)u_k^*-G(z_k)\left(\tilde{W}_{u,k}^T\varphi_u(z_k)+\varepsilon_u(z_k)\right),k=k_i. \tag{A.14}$$

With the above conclusions, now consider the Lyapunov function candidate given as

$$L_{cl,k}=L_{z,k}+L_{\tilde{z},k}+\Lambda_I L_{I,k}+\Lambda_V L_{V,k}+\Lambda_u L_{u,k}+\Lambda_{V2}L_{\tilde{W}_{V2},k}+\Lambda_{I2}L_{\tilde{W}_{I2},k}, \tag{A.15}$$

where $L_{z,k}=\underset{\tau,\gamma}{E}\{z_k^Tz_k\}$, $L_{\tilde{z},k}=\left\|\underset{\tau,\gamma}{E}\{\tilde{z}_k\}\right\|$, $L_{I,k}=tr\left(\underset{\tau,\gamma}{E}\{\tilde{W}_{I,k}^T\tilde{W}_{I,k}\}\right)$, $L_{V,k}=\underset{\tau,\gamma}{E}\{\tilde{W}_{V,k}^T\tilde{W}_{V,k}\}$, $L_{u,k}=tr\left(\underset{\tau,\gamma}{E}\{\tilde{W}_{u,k}^T\tilde{W}_{u,k}\}\right)$, $L_{\tilde{W}_{V2},k}=\left(\underset{\tau,\gamma}{E}\{\tilde{W}_{V,k}^T\tilde{W}_{V,k}\}\right)^2$ and $L_{\tilde{W}_{I2},k}=tr\left(\underset{\tau,\gamma}{E}\{\tilde{W}_{I,k}^T\tilde{W}_{I,k}\}\right)^2$. The positive constants coefficients are defined as

$$\Lambda_I=\frac{\Phi_{I,M}\varphi_{u,M}+4\Lambda_u(3\alpha_u+4\alpha_u^2)\lambda_{\max}^2(R_z^{-1})\phi_{V,M}'^2W_{V,M}^2\Phi_{I,M}^2}{\alpha_I(1-2\alpha_I)\bar{\Phi}_{I,m}^2},$$

$$\Lambda_V=3\Lambda_u(3\alpha_u+4\alpha_u^2)\phi_{V,M}'^2\lambda_{\max}^2(R_z^{-1})\Phi_{I,M}^2W_{I,M}^2\big/\alpha_V\left(1-2\alpha_V\right)\bar{\psi}_m,$$

$$\Lambda_u=2(4G_M^2\varphi_{u,M}^2+\Phi_{I,M}\varphi_{u,M})\big/\alpha_u(1-4\alpha_u)\bar{\varphi}_{u,m}$$

$$\Lambda_{V2}=\frac{\Lambda_u 2(3\alpha_u+4\alpha_u^2)\phi_{V,M}'^2\lambda_{\max}^2(R_z^{-1})\Phi_{I,M}^2}{(2-\alpha_V(1-2\alpha_V)\bar{\psi}_m)(\alpha_V(1-2\alpha_V)\bar{\psi}_m)},\text{ and}$$

$$\Lambda_{I2}=\frac{\Lambda_u(3\alpha_u+4\alpha_u^2)\phi_{V,M}'^2\lambda_{\max}^2(R_z^{-1})\Phi_{I,M}^2}{(2-\alpha_I(1-2\alpha_I)\bar{\Phi}_{I,m})\alpha_I(1-2\alpha_I)\bar{\Phi}_{I,m}}$$

We will consider each term in (A.15) individually and combine the individual first differences to compute the overall first difference. Consider the first term $L_{z,k} = \underset{\tau,\gamma}{E}\{z_k^T z_k\}$ of the Lyapunov function (A.15). The first difference along the closed-loop system dynamics (A.14) can be expressed as

$$\begin{aligned}\Delta L_{z,k} &= \underset{\tau,\gamma}{E}\{z_{k+1}^T z_{k+1}\} - \underset{\tau,\gamma}{E}\{z_k^T z_k\} \\ &= \underset{\tau,\gamma}{E}\Big\{(F(z_k)+G(z_k)u_k^* - G(z_k)(\tilde{W}_{u,k}^T\varphi_u(z_k)+\varepsilon_u(z_k)))^T \\ &(F(z_k)+G(z_k)u_k^* - G(z_k)(\tilde{W}_{u,k}^T\varphi_u(z_k)+\varepsilon_u(z_k)))\Big\} - \underset{\tau,\gamma}{E}\{z_k^T z_k\}.\end{aligned}$$

By using Frobenius norm and applying C-S inequality $\left(a_1+a_2+\cdots+a_n\right)^2 \le na_1^2+\cdots+na_n^2$, the first difference is bounded by

$$\Delta L_{z,k} \le 2\left\|\underset{\tau,\gamma}{E}\{F(z_k)+G(z_k)u_k^*\}\right\|^2 + 4G_M^2\varphi_{u,M}^2\left\|\underset{\tau,\gamma}{E}\{\tilde{W}_{u,k}\}\right\|^2 + 4G_M^2\varepsilon_{u,M}^2 - \left\|\underset{\tau,\gamma}{E}\{z_k\}\right\|^2.$$

Recalling the Lemma 2, the first difference becomes

$$\Delta L_{z,k} \le -(1-2\mu)\left\|\underset{\tau,\gamma}{E}\{z_k\}\right\|^2 + 4G_M^2\varphi_{u,M}^2\left\|\underset{\tau,\gamma}{E}\{\tilde{W}_{u,k}\}\right\|^2 + 4G_M^2\varepsilon_{u,M}^2. \tag{A.16}$$

Moving on for the second term $L_{\tilde{z},k} = \left\|\underset{\tau,\gamma}{E}\{\tilde{z}_k\}\right\|$ in (A.15), , the first difference along the identification error dynamics (A.2) can be computed as

$$\Delta L_{\tilde{z},k} = \left\|\underset{\tau,\gamma}{E}\{\tilde{z}_{k+1}\}\right\| - \left\|\underset{\tau,\gamma}{E}\{\tilde{z}_k\}\right\| = \left\|\underset{\tau,\gamma}{E}\{\tilde{W}_{I,k}^T\Phi_I(z_k)\bar{u}_k + \bar{\varepsilon}_{I,k}\}\right\| - \left\|\underset{\tau,\gamma}{E}\{\tilde{z}_k\}\right\|.$$

Substituting the control input (34) and applying C-S inequality, the first difference leads to

$$\begin{aligned}\Delta L_{\tilde{z},k} \le &-\left\|\underset{\tau,\gamma}{E}\{\tilde{z}_k\}\right\| + (1/2)\Phi_{I,M}\varphi_{u,M}\left\|\underset{\tau,\gamma}{E}\{\tilde{W}_{I,k}\}\right\|^2 + (1/2)\Phi_{I,M}\varphi_{u,M}\left\|\hat{W}_{u,k}^T\varphi_u(z_k)\right\|^2 \\ &+(1/2)\Phi_{I,M}\varphi_{u,M} + \left\|\underset{\tau,\gamma}{E}\{\bar{\varepsilon}_{I,k}\}\right\|.\end{aligned} \tag{A.17}$$

Next, considering the third term , $L_{I,k} = tr\left(\underset{\tau,\gamma}{E}\{\tilde{W}_{I,k}^T \tilde{W}_{I,k}\}\right)$, the first difference remains same as in (A.4) and given by

$$\Delta L_{I,k} \le -\bar{\Phi}_{I,m}^2 \alpha_I (1-2\alpha_I) \left\| \underset{\tau,\gamma}{E}\{\tilde{W}_{I,k}\} \right\|^2 + (1+2\alpha_I)\alpha_I \bar{\varepsilon}_{I,M}^2 . \tag{A.18}$$

Now, considering fourth the term, $L_{V,k} = \underset{\tau,\gamma}{E}\{\tilde{W}_{V,k}^T \tilde{W}_{V,k}\}$, the first difference is same as in (A.10) of Lemma 1 and given by

$$\Delta L_{V,k} \le -\alpha_V (1-2\alpha_V)\bar{\psi}_m \left\| \underset{\tau,\gamma}{E}\{\tilde{W}_{V,k}\} \right\|^2 + (1+2\alpha_V)\alpha_V \Xi_{V,M}^2 . \tag{A.19}$$

Considering the fifth term, $L_{u,k} = tr\left(\underset{\tau,\gamma}{E}\{\tilde{W}_{u,k}^T \tilde{W}_{u,k}\}\right)$, the first difference is given by

$$\Delta L_{u,k} = tr\left(\underset{\tau,\gamma}{E}\{\tilde{W}_{u,k+1}^T \tilde{W}_{u,k+1}\}\right) - tr\left(\underset{\tau,\gamma}{E}\{\tilde{W}_{u,k}^T \tilde{W}_{u,k}\}\right)$$

$$= 2\alpha_u tr\left(\underset{\tau,\gamma}{E}\left\{ \frac{\tilde{W}_{u,k}^T \varphi_u(z_k) e_{u,k}^T}{\varphi_u^T(z_k)\varphi_u(z_k)+1} \right\} \right) + \alpha_u^2 tr\left(\underset{\tau,\gamma}{E}\left\{ \frac{e_{u,k}\varphi_u^T(z_k)\varphi_u(z_k)e_{u,k}^T}{\left(\varphi_u^T(z_k)\varphi_u(z_k)+1\right)^2} \right\} \right) \tag{A.20}$$

Substituting the control input error , $e_{u,k}$, from (A.12), applying the C-S and replacing with upper bounds i.e., $\frac{\varphi_u(z_k)\varphi_u^T(z_k)}{\varphi_u^T(z_k)\varphi_u(z_k)+1} \le 1$, the first difference is bounded by

$$\Delta L_{u,k} \le -\alpha_u tr\left(\underset{\tau,\gamma}{E}\left\{ \frac{\tilde{W}_{u,k}^T \varphi_u(z_k)\varphi_u^T(z_k)\tilde{W}_{u,k}}{\varphi_u^T(z_k)\varphi_u(z_k)+1} \right\} \right) + 4\alpha_u^2 tr\left(\underset{\tau,\gamma}{E}\left\{ \frac{(\tilde{W}_u^T\varphi_u(z_k))(\tilde{W}_u^T\varphi_u(z_k))^T}{\varphi_u^T(z_k)\varphi_u(z_k)+1} \right\} \right)$$
$$+(3\alpha_u + 4\alpha_u^2) tr\left\{ \underset{\tau,\gamma}{E}\left(\left(R_z^{-1}\tilde{G}^T(z_k)\left(\partial\phi_V(z_{k+1})/\partial z_{k+1}\right)W_V \right)\left(R_z^{-1}\tilde{G}^T(z_k)\left(\partial\phi_V(z_{k+1})/\partial z_{k+1}\right)W_V \right)^T \right) \right\}$$
$$+3(3\alpha_u + 4\alpha_u^2) tr\left(\underset{\tau,\gamma}{E}\left\{ \left(R_z^{-1}\hat{G}^T(z_k)\left(\partial\phi_V(z_{k+1})/\partial z_{k+1}\right)\tilde{W}_{V,k} \right)^T \left(R_z^{-1}\hat{G}^T(z_k)\left(\partial\phi_V(z_{k+1})/\partial z_{k+1}\right)\tilde{W}_{V,k} \right) \right\} \right)$$
$$+3\alpha_u tr\left(\underset{\tau,\gamma}{E}\{\varepsilon_{e,k}\varepsilon_{e,k}^T\} \right) + 4\alpha_u^2 tr\left(\underset{\tau,\gamma}{E}\{\varepsilon_{e,k}\varepsilon_{e,k}^T\} \right)$$

By definition $\tilde{W}_{I,k} = W_I - \hat{W}_{I,k}$ and using the identifier dynamics for $k = k_i$, we have $\left\| \underset{\tau,\gamma}{E}\{\hat{G}(z_k)\} \right\| \le \left\| \underset{\tau,\gamma}{E}\{\hat{W}_{I,k}^T \Phi_I(z_k)\} \right\|$ and $\left\| \underset{\tau,\gamma}{E}\{\tilde{G}(z_k)\} \right\|^2 \le \left\| \underset{\tau,\gamma}{E}\{\tilde{W}_{I,k}^T \Phi_I(z_k) + \varepsilon_I(z_k)\} \right\|$.

Taking the Frobenius norm and a using the above inequalities, the first difference satisfies the following inequality

$$\begin{aligned}\Delta L_{u,k} \le & -\alpha_u(1-4\alpha_u)\bar{\varphi}_{u,m}\left\|\underset{\tau,\gamma}{E}\{\tilde{W}_{u,k}\}\right\|^2 + 2(3\alpha_u+4\alpha_u^2)\lambda_{\max}^2(R_z^{-1})\phi_{V,M}^{'2}W_{V,M}^2\Phi_{I,M}^2 \\ & \times\left\|\underset{\tau,\gamma}{E}\{\tilde{W}_{I,k}\}\right\|^2 + 2(3\alpha_u+4\alpha_u^2)\phi_{V,M}^{'2}\lambda_{\max}^2(R_z^{-1})\Phi_{I,M}^2W_{I,M}^2\left\|\underset{\tau,\gamma}{E}\{\tilde{W}_{V,k}\}\right\|^2 \\ & +(1/2)(3\alpha_u+4\alpha_u^2)\phi_{V,M}^{'2}\lambda_{\max}^2(R_z^{-1})\Phi_{I,M}^2\left\|\underset{\tau,\gamma}{E}\{\hat{W}_{I,k}\}\right\|^4 \\ & +(1/2)(3\alpha_u+4\alpha_u^2)\phi_{V,M}^{'2}\lambda_{\max}^2(R_z^{-1})\Phi_{I,M}^2\left\|\underset{\tau,\gamma}{E}\{\tilde{W}_{V,k}\}\right\|^4 + \Xi_u^M,\end{aligned} \tag{A.21}$$

where $\Xi_u^M = 3\alpha_u\varepsilon_{I.M}^2 + 4\alpha_u^2\varepsilon_{I.M}^2 + 2(3\alpha_u+4\alpha_u^2)\lambda_{\max}^2(R_z^{-1})\phi_{V,M}^{'2}W_{V,M}^2\varepsilon_{I.M}^2$

Considering the sixth term, $L_{\tilde{W}_{V2},k} = \left(\underset{\tau,\gamma}{E}\{\tilde{W}_{V,k}^T\tilde{W}_{V,k}\}\right)^2$, the first difference can be written as

$$\Delta L_{\tilde{W}_{V2},k} = \left(\underset{\tau,\gamma}{E}\{\tilde{W}_{V,k+1}^T\tilde{W}_{V,k+1}\}\right)^2 - \left(\underset{\tau,\gamma}{E}\{\tilde{W}_{V,k}^T\tilde{W}_{V,k}\}\right)^2$$

$$= \left(\underset{\tau,\gamma}{E}\{\tilde{W}_{V,k+1}^T\tilde{W}_{V,k+1}\} - \underset{\tau,\gamma}{E}\{\tilde{W}_{V,k}^T\tilde{W}_{V,k}\} + 2\underset{\tau,\gamma}{E}\{\tilde{W}_{V,k}^T\tilde{W}_{V,k}\}\right)\left(\underset{\tau,\gamma}{E}\{\tilde{W}_{V,k+1}^T\tilde{W}_{V,k+1}\} - \underset{\tau,\gamma}{E}\{\tilde{W}_{V,k}^T\tilde{W}_{V,k}\}\right).$$

Substituting the first difference from (A.19), and with simple mathematical operations, the first difference is bounded by

$$\begin{aligned}\Delta L_{\tilde{W}_{V2},k} \le & -(1/2)\left(2-\alpha_V(1-2\alpha_V)\bar{\psi}_m\right)\left(\alpha_V(1-2\alpha_V)\bar{\psi}_m\right)\left\|\underset{\tau,\gamma}{E}\{\tilde{W}_{V,k}\}\right\|^4 \\ & + \frac{(1+2\alpha_V)^2\alpha_V^2}{2\alpha_V\bar{\psi}_m(1-2\alpha_V)}\left(2-\alpha_V(1-2\alpha_V)\bar{\psi}_m\right)\Xi_{V,M}^4 + (1+2\alpha_V)^2\alpha_V^2\Xi_{V,M}^4\end{aligned} \tag{A.22}$$

Considering the last term, $V_{\tilde{W}_{I2},k} = tr\left(\underset{\tau,\gamma}{E}\{\tilde{W}_{I,k}^T\tilde{W}_{I,k}\}\right)^2$, the first difference using the first difference (A.18) , similar to the previous term, computed as

$$\Delta L_{\tilde{W}_{I2},k} \le -(1/2)(2-\alpha_I(1-2\alpha_I)\bar{\Phi}_{I,m})\alpha_I(1-2\alpha_I)\bar{\Phi}_{I,m}$$
$$\left\|\underset{\tau,\gamma}{E}\{\tilde{W}_{I,k}\}\right\|^4 + \frac{1}{2}\frac{(2\alpha_I^2+\alpha_I)^2}{\alpha_I(1-2\alpha_I)\bar{\Phi}_{I,m}}\left(2-\alpha_I(1-2\alpha_I)\bar{\Phi}_{I,m}\right)\bar{\varepsilon}_{I,M}^4 \quad \text{(A.23)}$$
$$+(2\alpha_I^2+\alpha_I)\bar{\varepsilon}_{I,M}^4$$

Finally, combining all the individual first differences (A.16), (A.17), (A.18), (A.19), (A.21), (A.22), and (A.23), and recalling the definition of $\Lambda_I, \Lambda_V, \Lambda_u, \Lambda_{V2}$, and Λ_{I2},the overall first difference of the Lyapunov function is given by

$$\Delta L_{cl,k} \le -(1-2\mu)\left\|\underset{\tau,\gamma}{E}\{z_k\}\right\|^2 - \left\|\underset{\tau,\gamma}{E}\{\tilde{z}_k\}\right\| - \left((1/2)\Phi_{I,M}\varphi_{u,M} + 2\Lambda_u(3\alpha_u+4\alpha_u^2)\right.$$
$$\left.\lambda_{\max}^2(R_z^{-1})\phi_{V,M}^{'2}W_{V,M}^2\Phi_{I,M}^2\right)\left\|\underset{\tau,\gamma}{E}\{\tilde{W}_{I,k}\}\right\|^2 - \Lambda_u(3\alpha_u+4\alpha_u^2)\phi_{V,M}^{'2}\lambda_{\max}^2(R_z^{-1})\Phi_{I,M}^2W_{I,M}^2$$
$$\times\left\|\underset{\tau,\gamma}{E}\{\tilde{W}_{V,k}\}\right\|^2 - \left(4G_M^2\varphi_{u,M}^2 + (1/2)\Phi_{I,M}\varphi_{u,M}\right)\left\|\underset{\tau,\gamma}{E}\{\tilde{W}_{u,k}\}\right\|^2 - (1/2)\Lambda_u(3\alpha_u+4\alpha_u^2) \quad \text{(A.24)}$$
$$\times\phi_{V,M}^{'2}\lambda_{\max}^2(R_z^{-1})\Phi_{I,M}^2\left\|\underset{\tau,\gamma}{E}\{\tilde{W}_{V,k}\}\right\|^4 - (1/2)\Lambda_u(3\alpha_u+4\alpha_u^2)\phi_{V,M}^{'2}\lambda_{\max}^2(R_z^{-1})$$
$$\times\Phi_{I,M}^2\left\|\underset{\tau,\gamma}{E}\{\tilde{W}_{I,k}\}\right\|^4 + \Xi_{total}^{cl},$$

where

$$\Xi_{total}^{cl} = \Lambda_{I2}\frac{(2\alpha_I^2+\alpha_I)^2}{2\alpha_I(1-2\alpha_I)\bar{\Phi}_{I,m}}(2-\alpha_I(1-2\alpha_I)\bar{\Phi}_{I,m})\bar{\varepsilon}_{I.M}^4 + \Lambda_{I2}(2\alpha_I^2+\alpha_I)\bar{\varepsilon}_{I.M}^4$$
$$+\Lambda_{V2}\frac{(1+2\alpha_V)^2\alpha_V^2}{2\alpha_V\bar{\psi}_m(1-2\alpha_V)}\left(2-\alpha_V(1-2\alpha_V)\bar{\psi}_m\right)\Xi_{V,M}^4$$
$$+\Xi_u^M + \Lambda_{V2}(1+2\alpha_V)^2\alpha_V^2\Xi_{V,M}^4 + 2\Lambda_I\alpha_I^2\bar{\varepsilon}_{I.M}^2 + \Lambda_I\alpha_I\bar{\varepsilon}_{I.M}^2 + 4G_M^2\varepsilon_{u,M}^2$$
$$+\bar{\varepsilon}_{I.M} + \Lambda_V(1+2\alpha_V)\alpha_V\Xi_{V,M}^2 + \Phi_{I,M}\varphi_{u,M}W_{u,M}^2 + \frac{1}{2}\Phi_{I,M}\varphi_{u,M}.$$

Define the following new variables for simplicity.

$$\vartheta_1 = \left((1/2)\Phi_{I,M}\varphi_{u,M} + 4\Lambda_u(3\alpha_u+4\alpha_u^2)\lambda_{\max}^2(R_z^{-1})\phi_{V,M}^{'2}W_{V,M}^2\Phi_{I,M}^2\right),$$

$$\vartheta_2 = (1/2)\Lambda_u(3\alpha_u+4\alpha_u^2)\phi_{V,M}^{'2}\lambda_{\max}^2(R_z^{-1})\Phi_{I,M}^2,$$

$$\varpi_1 = \Lambda_u(3\alpha_u+4\alpha_u^2)\phi_{V,M}^{'2}\lambda_{\max}^2(R_z^{-1})\Phi_{I,M}^2W_{I,M}^2, \text{ and}$$

$$\varpi_2 = (1/2)\Lambda_u(3\alpha_u + 4\alpha_u^2)\phi_{V,M}^{'2}\lambda_{\max}^2(R_z^{-1})\Phi_{I,M}^2 .$$

From (A.24), the overall first difference $\Delta L_{cl,k} < 0$ as long as

$$\left\| \underset{\tau,\gamma}{E}\{z_k\} \right\| > \sqrt{\Xi_{total}^{cl}/(1-2\mu)} \equiv B_z^1, \text{ or } \left\| \underset{\tau,\gamma}{E}\{\tilde{z}_k\} \right\| > \sqrt{\Xi_{total}^{cl}} \text{ or}$$

$$\left\| \underset{\tau,\gamma}{E}\{\tilde{W}_{I,k}\} \right\| > \max\left(\sqrt{\Xi_{total}^{cl}/\vartheta_1}, \sqrt[4]{\Xi_{total}^{cl}/\vartheta_2} \right) \equiv B_{\tilde{W}_I}^1, \quad \text{or}$$

$$\left\| \underset{\tau,\gamma}{E}\{\tilde{W}_{V,k}\} \right\| > \max\left(\sqrt{\Xi_{total}^{cl}/\varpi_1}, \sqrt[4]{2\Xi_{total}^{cl}/\varpi_2} \right) \equiv B_{\tilde{W}_V}^1 \quad \text{or}$$

$$\left\| \underset{\tau,\gamma}{E}\{\tilde{W}_{u,k}\} \right\| > \sqrt{2\Xi_{total}^{cl}/(8G_M^2\varphi_{u,M}^2 + \Phi_{I,M}\varphi_{u,M})} \equiv B_{\tilde{W}_u}^1 .$$

Therefore, $\underset{\tau,\gamma}{E}\{z_k\}$, $\underset{\tau,\gamma}{E}\{\tilde{z}_k\}$, $\underset{\tau,\gamma}{E}\{\tilde{W}_{I,k}\}$, $\underset{\tau,\gamma}{E}\{\tilde{W}_{V,k}\}$ and $\underset{\tau,\gamma}{E}\{\tilde{W}_{u,k}\}$ are UB in the mean at the event sampled instants.

Case II: During inter sample times ($k_i < k < k_{i+1}, \forall 1,2,\cdots$)

The closed loop system dynamics remains same as in (40) for $k_i < k < k_{i+1}$. Now consider a Lyapunov function candidate same as in Case I. The first difference of the first term along the closed-loop event-triggered system dynamics (40) for $k_i < k < k_{i+1}$ can be expressed as

$$\Delta L_{z,k} = \underset{\tau,\gamma}{E}\left\{ \left(F(z_k) + G(z_k)u_k^* - G(z_k)\hat{W}_{u,k}^T\left(\varphi_u(z_k) - \varphi_u(\breve{z}_k)\right) - G(z_k)\left(\tilde{W}_{u,k}^T\varphi_u(z_k) + \varepsilon_u(z_k)\right)\right)^T \right.$$
$$\left. \times\left(F(z_k) + G(z_k)u_k^* - G(z_k)\hat{W}_{u,k}^T\left(\varphi_u(z_k) - \varphi_u(\breve{z}_k)\right) - G(z_k)\left(\tilde{W}_{u,k}^T\varphi_u(z_k) + \varepsilon_u(z_k)\right)\right) \right\} - \underset{\tau,\gamma}{E}\{z_k^T z_k\}$$

Recall the Lipschitz continuity of the actor NN activation function from Assumption 5. Using Frobenius norm and C-S inequality, the first difference satisfies

$$\Delta L_{z,k} \le -(1-2\mu)\left\| \underset{\tau,\gamma}{E}\{z_k\} \right\|^2 + 6G_M^2\varphi_{u,M}^2\left\| \underset{\tau,\gamma}{E}\{\tilde{W}_{u,k}\} \right\|^2 + 6G_M^2C_{\varphi_u}^2e_{ET,k}^2\left\| \underset{\tau,\gamma}{E}\{\hat{W}_{u,k}\} \right\|^2 \quad \text{(A.25)}$$
$$+6G_M^2\varepsilon_{u,M}^2$$

Considering the second term, the first difference along the identification error dynamics (20) for $k_i < k < k_{i+1}$

$$\begin{aligned}\Delta L_{\tilde{z},k} \le &-\left\|\underset{\tau,\gamma}{E}\{\tilde{z}_k\}\right\| + \Phi_{I,M}\left\|\underset{\tau,\gamma}{E}\{\tilde{W}_{I,k}\}\right\|\left\|\underset{\tau,\gamma}{E}\{\bar{u}_k\}\right\| + \left\|\underset{\tau,\gamma}{E}\{\bar{\varepsilon}_{I,k}\}\right\| \\ &+\left\|\underset{\tau,\gamma}{E}\{\hat{W}_{I,k}^T\left(\Phi_I(z_k)-\Phi_I(\breve{z}_k)\right)\bar{u}_k\}\right\|\end{aligned} \tag{A.26}$$

Since the identifier, critic and actor NN weights are not updated and held at their previous values, the first difference of the rest terms are zero, i.e.,

$$\Delta L_{I,k} = 0,\ \Delta L_{V,k} = 0,\ \Delta L_{u,k} = 0,\ \Delta L_{\tilde{W}_{V2},k} = 0,\ \Delta L_{\tilde{W}_{I2},k} = 0 \tag{A.27}$$

Finally combining all individual first differences (A.25), (A.26), and (A.27), the total first difference of the Lyapunov function becomes

$$\begin{aligned}\Delta L_{cl,k} \le &-(1-2\mu)\left\|\underset{\tau,\gamma}{E}\{z_k\}\right\|^2 - \left\|\underset{\tau,\gamma}{E}\{\tilde{z}_k\}\right\| + 6G_M^2 C_{\varphi_u}^2 e_{ET,k}^2\left\|\underset{\tau,\gamma}{E}\{\hat{W}_{u,k}\}\right\|^2 + (1/2)C_{\Phi_I} \\ &\times\left\|\underset{\tau,\gamma}{E}\{\hat{W}_{I,k}^T\}\right\|^2 e_{ET,k}^2 + 6G_M^2\varphi_{u,M}^2\left\|\underset{\tau,\gamma}{E}\{\tilde{W}_{u,k}\}\right\|^2 + (1/2)\Phi_{I,M}\left\|\underset{\tau,\gamma}{E}\{\tilde{W}_{I,k}\}\right\|^2 \\ &+(1/2)(\Phi_{I,M}+C_{\Phi_I})\left\|\underset{\tau,\gamma}{E}\{\bar{u}_k\}\right\|^2 + 6G_M^2\varepsilon_{u,M}^2 + \left\|\underset{\tau,\gamma}{E}\{\bar{\varepsilon}_{I,k}\}\right\|\end{aligned}$$

Recalling the event-trigger condition(42) , $\hat{W}_I = W_I - \tilde{W}_I$ and applying C-S inequality, the first difference satisfies

$$\Delta L_{cl,k} \le -(1-2\mu)(1-\Gamma)\left\|\underset{\tau,\gamma}{E}\{z_k\}\right\|^2 - \left\|\underset{\tau,\gamma}{E}\{\tilde{z}_k\}\right\| + B_{\tilde{W},k}^{c2} \tag{A.28}$$

where

$$\begin{aligned}B_{\tilde{W},k}^{c2} = &\,6G_M^2\varphi_{u,M}^2\left\|\underset{\tau,\gamma}{E}\{\tilde{W}_{u,k}\}\right\|^2 + (1/2)\Phi_{I,M}\left\|\underset{\tau,\gamma}{E}\{\tilde{W}_{I,k}\}\right\|^2 + (\Phi_{I,M}+C_{\Phi_I})\varphi_{u,M}\left\|\underset{\tau,\gamma}{E}\{\tilde{W}_{u,k}\}\right\|^2 \\ &+(1/2)(\Phi_{I,M}+C_{\Phi_I}) + \Phi_{I,M}\varphi_{u,M}W_{u,M}^2 + \bar{\varepsilon}_{I,M} + 6G_M^2\varepsilon_{u,M}^2.\end{aligned}$$

From (A.28), it is evident that the first difference of the Lyapunov function is less than zero as long as

$$\left\| \underset{\tau,\gamma}{E}\{z_k\} \right\| > \sqrt{B_{\tilde{W},k}^{c2} / \left(1-2\mu\right)\left(1-\Gamma\right)} \equiv B_{z,k}^{c2}, k_i < k < k_{i+1} \text{ or}$$

$$\left\| \underset{\tau,\gamma}{E}\{z_k\} \right\| > \sqrt{B_{\tilde{W},k}^{c2}} \equiv B_{\tilde{z},k}^{c2}, k_i < k < k_{i+1}.$$

This implies $\underset{\tau,\gamma}{E}\{z_k\}$, $\underset{\tau,\gamma}{E}\{\tilde{z}_k\}$, $k_i < k < k_{i+1}$ are bounded. Note that the initial NN weight estimation errors are bounded due to finite initial values and bounded target weights. Again, from(A.27), the NN weight estimation errors $\underset{\tau,\gamma}{E}\{\tilde{W}_{I,k_i+1}\} = \underset{\tau,\gamma}{E}\{\tilde{W}_{I,k}\} = \underset{\tau,\gamma}{E}\{\tilde{W}_{I,k_{i+1}}\}$, $\underset{\tau,\gamma}{E}\{\tilde{W}_{V,k_i+1}\} = \underset{\tau,\gamma}{E}\{\tilde{W}_{V,k}\} = \underset{\tau,\gamma}{E}\{\tilde{W}_{V,k_{i+1}}\}$ and $\underset{\tau,\gamma}{E}\{\tilde{W}_{u,k_i+1}\} = \underset{\tau,\gamma}{E}\{\tilde{W}_{u,k}\} = \underset{\tau,\gamma}{E}\{\tilde{W}_{u,k_{i+1}}\}$ for $k_i < k < k_{i+1}$, $\forall 1,2,\cdots$. Further, from Case I, $\underset{\tau,\gamma}{E}\{\tilde{W}_{I,k}\}$, $\underset{\tau,\gamma}{E}\{\tilde{W}_{V,k}\}$ and $\underset{\tau,\gamma}{E}\{\tilde{W}_{u,k}\}$, $k = k_i, \forall 1,2,\cdots$ are bounded in the mean. Therefore, $\underset{\tau,\gamma}{E}\{\tilde{W}_{I,k}\}$, $\underset{\tau,\gamma}{E}\{\tilde{W}_{V,k}\}$ and $\underset{\tau,\gamma}{E}\{\tilde{W}_{u,k}\}, k_i < k < k_{i+1}$ are constant and bounded during the inter-sample times.

Further, from Case II, $\underset{\tau,\gamma}{E}\{\tilde{W}_{I,k_i+1}\} = \underset{\tau,\gamma}{E}\{\tilde{W}_{I,k}\} = \underset{\tau,\gamma}{E}\{\tilde{W}_{I,k_{i+1}}\}$ and $\underset{\tau,\gamma}{E}\{\tilde{W}_{u,k_i+1}\} = \underset{\tau,\gamma}{E}\{\tilde{W}_{u,k}\} = \underset{\tau,\gamma}{E}\{\tilde{W}_{u,k_{i+1}}\}$, $k_i < k < k_{i+1}$, $\forall 1,2,\cdots$. Thus, $B_{\tilde{W},k}^{c2}$, $k_i < k < k_{i+1}$ is a piece wise constant function. This follows that $B_{z,k}^{c2}$ and $B_{\tilde{z},k}^{c2}$, $k_i < k < k_{i+1}$ in (A.28) are also piece wise constant functions

Next, we will combine both the cases to show the convergence of closed-loop parameters $\underset{\tau,\gamma}{E}\{z_k\}$, $\underset{\tau,\gamma}{E}\{\tilde{z}_k\}$ $\underset{\tau,\gamma}{E}\{\tilde{W}_{I,k}\}$, $\underset{\tau,\gamma}{E}\{\tilde{W}_{V,k}\}$ and $\underset{\tau,\gamma}{E}\{\tilde{W}_{u,k}\}$ to the UB in the mean. From Case I, $\underset{\tau,\gamma}{E}\{\tilde{W}_{I,k_i+1}\} \le \underset{\tau,\gamma}{E}\{\tilde{W}_{I,k_i}\}$ and $\underset{\tau,\gamma}{E}\{\tilde{W}_{u,k_i+1}\} \le \underset{\tau,\gamma}{E}\{\tilde{W}_{u,k_i}\}$ and from Case II, $\underset{\tau,\gamma}{E}\{\tilde{W}_{I,k_i+1}\} = \underset{\tau,\gamma}{E}\{\tilde{W}_{I,k_{i+1}}\}$ and $\underset{\tau,\gamma}{E}\{\tilde{W}_{u,k_i+1}\} = \underset{\tau,\gamma}{E}\{\tilde{W}_{u,k_{i+1}}\}$, $\forall 1,2,\cdots$. Therefore, it holds that $\underset{\tau,\gamma}{E}\{\tilde{W}_{I,k_{i+1}}\} \le \underset{\tau,\gamma}{E}\{\tilde{W}_{I,k_i}\}$ and $\underset{\tau,\gamma}{E}\{\tilde{W}_{u,k_{i+1}}\} \le \underset{\tau,\gamma}{E}\{\tilde{W}_{u,k_i}\}$. This implies the bound $B_{\tilde{W},k_{i+1}}^{c2} \le B_{\tilde{W},k_i}^{c2}$,

$\forall 1,2,\cdots$. This further implies $B_{z,k_{i+1}}^{c2} \le B_{z,k_i}^{c2}$ and $B_{z,k_{i+1}}^{c2} \le B_{z,k_i}^{c2}$. It follows that the bounds $B_{z,k_i}^{c2} \to B_{z,M}^{c2}$ and $B_{\tilde{z},k_i}^{c2} \to B_{\tilde{z},M}^{c2}$ for all $k_i > k_0 + N$ where $B_{z,M}^{c2} = \sqrt{B_{\tilde{W},M}^{c2} \big/ (1-2\mu)(1-\Gamma)}$ and $B_{\tilde{z},M}^{c2} = \sqrt{B_{\tilde{W},M}^{c2}}$ with $B_{\tilde{W},M}^{c2}$ computed form $B_{\tilde{W},k}^{c2}$ in (A.28) by replacing $\underset{\tau,\gamma}{E}\{\tilde{W}_{I,k}\}$ and $\underset{\tau,\gamma}{E}\{\tilde{W}_{u,k}\}$ with its UB $B_{\tilde{W}_I}^1$ and $B_{\tilde{W}_u}^1$, respectively, from Case I.

Consequently, the closed-loop parameters $\underset{\tau,\gamma}{E}\{z_k\}$, $\underset{\tau,\gamma}{E}\{\tilde{z}_k\}$, $\underset{\tau,\gamma}{E}\{\tilde{W}_{I,k}\}$, $\underset{\tau,\gamma}{E}\{\tilde{W}_{V,k}\}$ and $\underset{\tau,\gamma}{E}\{\tilde{W}_{u,k}\}$ are UB in the mean for all sampling instants $k_i > k_0 + N$ or alternatively, $k > k_0 + \bar{N}$ for $\bar{N} > N$ with ultimate bounds given by $B_z = \max\left(B_z^1, B_{z,M}^{c2}\right)$, $B_{\tilde{z}} = \max\left(B_{\tilde{z}}^1, B_{\tilde{z},M}^{c2}\right)$, $B_{\tilde{W}_I}^1$, $B_{\tilde{W}_V}^1$, and $B_{\tilde{W}_u}^1$, respectively..

Now to show the convergence of the estimated value function and control input to their respective optimal values consider the differences

$$\left\|\hat{V}_k - V_k^*\right\| = \left\|\hat{W}_{V,k}^T \phi_V(\breve{z}_k) - W_V^T \phi_V(\breve{z}_k) - \varepsilon_{e,V}(\breve{z}_k, e_{ET,k})\right\| \le \left\|\tilde{W}_V^T \phi_V(\breve{z}_k)\right\| + \left\|\varepsilon_{e,V}(\breve{z}_k, e_{ET,k})\right\|$$
$$\le \phi_{V,M} B_{\tilde{W}_V}^1 + W_{V,M} C_{\phi_V} \sigma_{ET,M} B_z + \varepsilon_{V,M} \equiv B_V \quad \text{and}$$

$$\left\|u_k - u_k^*\right\| = \left\|\hat{W}_{u,k}^T \varphi_u(\breve{z}_k) - W_u^T \varphi_u(\breve{z}_k) - \varepsilon_{e,u}(\breve{z}_k, e_{ET,k})\right\| \le \phi_{\mathrm{u},M} B_{\tilde{W}_V}^1 + W_{u,M} C_{\phi_u} \sigma_{ET,M} B_z + \varepsilon_{\mathrm{u},M} \equiv B_u$$

where B_V and B_u are small positive constants. ■

REFERENCES

[1] B. C. Kuo, *Analysis and Synthesis of Sampled Data Control Systems.* Englewood Cliffs, NJ: Prentice Hall, 2012.

[2] K**.** Ogata, *Discrete Time Control Systems*. Englewood Cliffs, NJ: Prentice Hall, 2011.

[3] K. Astrom and B. Bernhardsson, "Comparison of Riemann and Lebesgue sampling for first order stochastic systems*,"* in *Proceedings of the 41st IEEE Conference on Decision Control*, Las Vegas, Nevada, USA, vol. 2, Dec. 2002, pp. 2011–2016.

[4] L. W. Liou and A. Ray, "A stochastic regulator for integrated communication and control systems: part I—formulation of control law," *ASME Journal of Dynamic Systems, Measurement, and Control*, vol. 113, no. 4, pp. 604–611, Dec. 1991.

[5] H. Shousong and Z. Qixin, "Stochastic optimal control and analysis of stability of networked control systems with long delay," *Automatica*, vol. 39, no. 11, pp. 1877-1884, Nov. 2003.

[6] G. C. Walsh, Ye Hong, and L. G. Bushnell, "Stability analysis of networked control systems," *IEEE Transactions on Control System Technology*, vol.10, no. 3, pp. 438-446, May 2002.

[7] M. C. F. Donkers, *Networked and event-triggered control systems*. PhD dissertation, Eindhoven University of Technology, Netherlands, 2011.

[8] P. Ellis, "Extension of phase plane analysis to quantized systems," *IRE Transactions on Automatic Control*, vol. 4, no. 2, pp. 43-54, Nov. 1959.

[9] R. Dorf, M. Farren, and C. Phillips, "Adaptive sampling frequency for sampled data control systems," *IRE Transactions on Automatic Control*, vol. 7, no. 1, pp. 38-47, Jan. 1962.

[10] A. M. phillips and Masayoshi Tomizuka, "Multi rate estimation and control under time-varying data sampling with applications to information storage devices," in *Proceeding of the American Control Conference,* Seattle, WA, USA, Jun. 1995, pp. 4152-4155.

[11] D. Hristu-Varsakelis and P. R. Kumar, "Interrupt-based feedback control over a shared communication medium," in *Proceedings of the 41st IEEE Conference on Decision and Control,* Las Vegas, Nevada, USA, Dec. 2002, pp. 3223-3228.

[12] M. Rabi and J. S. Baras, "Level-triggered control of scalar linear system," in *Proceedings of the Mediterranean Conference on Control and Automation*, Athens, Greece, Jul.2007, pp. 1-6.

[13] K. E. Arzen, "A simple event-based PID controller," In *Proceedings of the 14th World Congress of IFAC*, vol. 18, Beijing, China, Jul. 1999, pp. 423-428.

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

[15] M. Mazo and P. Tabuda, "Decentralized event-triggered control over wireless sensor/actuator networks," *IEEE Transaction on Automatic Control*, vol. 56, no. 10, pp. 2456-2461, Aug. 2011.

[16] W. Heemels, J. Sandee, and P. van den Bosch, "Analysis of event-driven controllers for linear systems," *International Journal of Control*, vol.81, no. 4, pp. 571–590, Apr. 2008.

[17] J. Lunze and D. Lehmann, "A state-feedback approach to event-based control," *Automatica*, vol. 46, no.1, pp. 211–215, Jan. 2010.

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

[19] W. P. M. H. Heemels, K. H. Johansson, and P Tabuada "An introduction to event-triggered and self-triggered control," in *Proceedings of the 51st IEEE Conference on Decision and Control*, Maul, Hawaii, USA, Dec. 2012, pp. 3270-3285.

[20] P. Tallapragada and N. Chopra, "Event-triggered decentralized dynamic output feedback control for LTI systems," in *Proceedings of 3rd IFAC Workshop on Distributed Estimation and Control in Networked Systems*, vol. 3, no.1, Santa Barbara, CA, USA, Sep. 2012, pp. 31-36.

[21] R. Cogill, "Event-based control using quadratic approximate value functions," in *Proceedings of the 48th Joint IEEE Conference on Decision and Control and 28th Chinese Control Conference*, Shanghai, P.R., China, Dec. 2009, pp. 5883–5888.

[22] X. Wang and M. Lemmon, "Event design in event-triggered feedback control systems," in *Proceedings of the 47th IEEE Conference on Decision Control*, Cancun, Mexico, Dec. 2008, pp. 2105–2110.

[23] 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 Transaction. on Automatic Control*, vol.57, no.6, pp.1362,1376, Jun. 2012.

[24] W. P. M. H. Heemels and M. C. F. Donkers, "Model-based periodic event-triggered control of linear systems," *Automatica*, vol. 49, no. 3, pp. 698-711, Mar. 2013.

[25] Eloy Garcia and Panos J. Antsaklis, "Parameter estimation in time-triggered and event-triggered model-based control of uncertain systems," *International Journal of Control*, vol. 85, no. 9, pp. 1327-1342, Apr. 2012.

[26] A. Eqtami, D. V. Dimarogonas, and K. J. Kyriakopoulos, "Event-triggered control for discrete-time systems," in *Proceeding of the American Control Conference*, Baltimore, MD, Jun. 2010, pp. 4719-4724.

[27] C. Stocker and J. Lunze, "Event-based control of nonlinear systems: An input-output linearization approach," in *Proceedings of the 50th Decision and Control and European Control Conference*, Orlando, FL, USA, Dec.2011, pp. 2541-2546.

[28] M. Mazo Jr. and M. Cao, "Decentralized event-triggered control with asynchronous updates," in *Proceedings of the 50th IEEE Conference on Decision and Control and European Control Conference*, Orlando, FL, USA, Dec. 2011, pp. 2547 -2552.

[29] M. Mazo Jr. and P. Tabuada, "Decentralized event-triggered control over Wireless Sensor/Actuator Networks," *IEEE Transactions on Automatic Control*, vol. 56, no. 10, pp. 2456-2461, Aug. 2011.

[30] X. Wang and M. Lemmon, "Event-triggering in distributed networked systems with data dropouts and delays," in Hybrid *systems: Computation and control*, R. Majumdar and P. Tabuada, Ed. Berlin Heidelberg, Germany: Springer-Verlag, pp. 366-380, 2009.

[31] X. Wang and M.D. Lemmon, "Event triggering in distributed networked control systems," *IEEE Transactions on Automatic Control*, vol. 56, no. 3, pp. 586-601, Mar. 2011.

[32] Tang Tao, Liu Zhongxin, and Chen Zengqiang, "Event-triggered formation control of multi-agent systems," in *Proceedings of the 30th Chinese Control Conference*, Yantai, China, Jul. 2011, pp. 4783-4786.

[33] D. V. Dimarogonas and K. H. Johansson, "Event-triggered control for multi agent systems," in *Proceedings of the 48th IEEE conference on Decision and Control and Chinese Control conference*, Shanghai, China, Dec. 2009, pp. 7131-7136.

[34] X. Wang, *Event-triggering in Cyber Physical System*. Ph.D. Dissertation, University of Notre Dame, USA, 2009.

[35] X. Wang and M.D. Lemmon, "Self-triggered feedback control systems with finite-gain L2 stability," *IEEE Transactions on Automatic Control*, vol. 54, no. 3, pp. 452-467, Mar. 2009.

[36] X. Wang and M.D. Lemmon, "Self-triggering under state-independent disturbances" *IEEE Transactions on Automatic Control*, vol. 55, no. 6, pp. 1494-1500, Jun. 2010.

[37] S. L. Hu and D. Yue, "Event-triggered control design of linear networked systems with quantization," *ISA Transactions*, vol. 51, no. 1, pp. 153–162, Jan. 2012.

[38] T. Henningsson, E. Johannesson, and A. Cervin, "Sporadic event-based control of first-order linear stochastic systems," *Automatica*, vol. 44, no. 11, pp. 2890-2895, Nov. 2008.

[39] D. Lehmann and J. Lunze, "Event-based output-feedback control," in *Proceedings of the 19th Mediterranean Conference on Control and Automatio*n, Corfu, Greece, Jun. 2011, pp. 982-987.

[40] A. Molin, S. Hirche, "On the optimality of certainty equivalence for event-triggered control systems," *IEEE Transactions on Automatic Control*, vol. 58, no. 2 , pp. 470-474, Feb. 2013.

[41] M. Rabi, K. H. Johansson, and M. Johansson, "Optimal stopping for event-triggered sensing and actuation," in *Proceedings of the 47th IEEE Conference on Decision and Control*, Cancun, Mexico, Dec. 2008, pp. 3607-3612.

[42] O.C. Imer and T. Basar, "To measure or to control: optimal control with scheduled measurements and controls," in *Proceedings of the American Control Conference*, Minneapolis, MN, USA, Jul. 2006, pp. 14-16.

[43] X. Wang and N. Hovakimyan, "$\mathcal{L}1$ adaptive control of event-triggered networked systems," in *Proceedings of the American Control Conference*, Baltimore, MD, Jun. 2010, pp. 2458-2463.

[44] E. Garcia and P. J. Antsaklis, "Model-based event-triggered control for systems with quantization and time-varying network delays," *IEEE Transaction on Automatic Control*, vol. 58, no. 2, pp. 422–434, Feb.2013.

[45] H. Xu, S. Jagannathan and F. L. Lewis, "Stochastic optimal control of unknown networked control systems in the presence of random delays and packet losses," *Automatica*, vol. 48, no. 6, pp. 1017-1030, Jun. 2012.

[46] M. Green and J. B. Moore, "Persistency of excitation in linear systems", *Systems and Control Letters*, vol. 7, pp. 351–360, Sep.1986.

[47] K. S. Narendra and A. M. Annaswamy, *Stable Adaptive Systems*. New Jersey: Prentice-Hall, 1989.

[48] F. L. Lewis and V. L. Syrmos, *Optimal Control.* 3rd edition. Hoboken, NJ: Wiley, 2012.

[49] H. K. Khalil, *Nonlinear Systems*. Upper Saddleriver, NJ: Prentice Hall, 2002.

[50] S.J. Bradtke, B.E. Ydestie, and A.G. Barto, "Adaptive linear quadratic control using policy iteration," in *Proceedings of the American Control Conference*, Baltimore, MD, USA, Jun.1994, pp. 3475–3476.

[51] C. J. C. H. Watkins, *Learning from Delayed Rewards*. PhD thesis, Cambridge University, Cambridge, England, 1989.

[52] R. Beard, *Improving the Closed-loop Performance of Nonlinear Systems*. Ph.D. dissertation, RPI, USA, 1995.

[53] D. P. Bertsekas and J. N. Tsitsiklis, *Neuro-Dynamic Programming.* MA: Athena Scientific, 1996.

[54] 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 Transaction on Neural Network and Learning Systems*, vol. 23, no. 7, pp. 1118-1129, Jul. 2012.

[55] Fei-Yue Wang, N. Jin, D. Liu, and Q. Wei, "Adaptive dynamic programming for finite-horizon optimal control of discrete-time nonlinear systems with ε -error bound" *IEEE Transactions on Neural Networks*, vol. 22, no. 1, pp. 24–36, Jan. 2011.

[56] G.C. Goodwin and K.S. Sin, *Adaptive Filtering Prediction and Control*, 1st edition. Mineola, New York: Dover publications, 1984.

[57] S. Hagen and B. Krose, "Linear quadratic regulation using reinforcement learning," in *Proceedings of the 8th Belgian-Dutch conference on machine learning*, 1998, pp. 39–46.

[58] S. Jagannathan, *Neural Network Control of Nonlinear Discrete-time Systems.* Boca Raton, FL: CRC Press, 2006.

[59] J. Si, A. Barto, W. Powell, and D. Wunsch, *Handbook of Learning and Approximate Dynamic Programming.* New Jersey: Wiley, 2004.

[60] Werbos, P. J. "Approximate dynamic programming for real-time control and neural modeling," in Handbook of intelligent control, D. A. White and D. A. Sofge, Eds. New York: Van Nostrand Reinhold, 1992.

[61] Q. Zhao, *Finite-Horizon Optimal Control of Linear and Class of Nonlinear Systems*. Ph. D. Thesis, Department of Electrical Engineering, Missouri University of Science and Technology, USA, 2013.

VITA

Avimanyu Sahoo was born on July 11, 1979 in Cuttack, Odisha, India. He earned the Bachelor of Science (B. S.) degree in Electrical Technology from Cochin University of Science and Technology, Cochin, Kerala, India and the Master of Technology (M. Tech.) degree in Electrical Engineering from Indian Institute of Technology (Banaras Hindu University) (IIT (BHU)) in 2008 and 2011, respectively. He joined Missouri University of Science and Technology (formerly the University of Missouri-Rolla) in August 2011 and received the degree of Doctor of Philosophy in Electrical Engineering in August 2015.