Fall 2020

Optimal adaptive control of time-delay dynamical systems with known and uncertain dynamics

Rohollah Moghadam

OPTIMAL ADAPTIVE CONTROL OF TIME-DELAY DYNAMICAL SYSTEMS WITH KNOWN AND UNCERTAIN DYNAMICS

by

ROHOLLAH MOGHADAM

A DISSERTATION

Presented to the Graduate Faculty of the

MISSOURI UNIVERSITY OF SCIENCE AND TECHNOLOGY

In Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

in

ELECTRICAL ENGINEERING

2020

Approved by

Dr. Jagannathan Sarangapani, Advisor
Dr. Kelvin Todd Erickson
Dr. K. Krishnamurthy
Dr. Maciej Zawodniok
Dr. Venkata Sriram Siddhardh Nadendla

PUBLICATION DISSERTATION OPTION

This dissertation consists of the following four articles which have been submitted or are under revision for publication as follows:

Paper I: Pages 8-33. A Linear Quadratic Regulator for Linear Time-invariant Systems with Input and State Delay, Revised and resubmitted to *International Journal of Control*.

Paper II: Pages 34-61. Optimal Adaptive Output Feedback Control of Uncertain Linear Continuous-time Systems with State and Input Delays, Revised and resubmitted to *Automatica*.

Paper III: Pages 62-98. Event-triggered Optimal Adaptive Control of Partially Unknown Linear Continuous-time Systems with State Delay, Under revision in *IEEE Transaction on System, Mans and Cybernetics*.

Paper IV: Pages 99-140. Optimal Adaptive Control of Uncertain Nonlinear Continuous-time Systems with Input and State Delays. Submitted to *IEEE Transaction on Neural Networks*.

ABSTRACT

Delays are found in many industrial pneumatic and hydraulic systems, and as a result, the performance of the overall closed-loop system deteriorates unless they are explicitly accounted. It is also possible that the dynamics of such systems are uncertain. On the other hand, optimal control of time-delay systems in the presence of known and uncertain dynamics by using state and output feedback is of paramount importance. Therefore, in this research, a suite of novel optimal adaptive control (OAC) techniques are undertaken for linear and nonlinear continuous time-delay systems in the presence of uncertain system dynamics using state and/or output feedback.

First, the optimal regulation of linear continuous-time systems with state and input delays by utilizing a quadratic cost function over infinite horizon is addressed using state and output feedback. Next, the optimal adaptive regulation is extended to uncertain linear continuous-time systems under a mild assumption that the bounds on system matrices are known. Subsequently, the event-triggered optimal adaptive regulation of partially unknown linear continuous time systems with state-delay is addressed by using integral reinforcement learning (IRL). It is demonstrated that the optimal control policy renders asymptotic stability of the closed-loop system provided the linear time-delayed system is controllable and observable. The proposed event-triggered approach relaxed the need for continuous availability of state vector and proven to be zeno-free.

Finally, the OAC using IRL neural network based control of uncertain nonlinear time-delay systems with input and state delays is investigated. An identifier is proposed for nonlinear time-delay systems to approximate the system dynamics and relax the need for the control coefficient matrix in generating the control policy. Lyapunov analysis is utilized to design the optimal adaptive controller, derive parameter/weight tuning law and verify stability of the closed-loop system.

ACKNOWLEDGMENTS

First, I would like to express my sincere gratitude to my advisor Prof. Jagannathan Sarangapani for the continuous support of my Ph.D study, and for his patience, motivation, and immense knowledge. His guidance helped me during the research and writing of my dissertation. It has been a great experience to work under his guidance, and I attribute the level of my PhD degree to his encouragement and effort. I will forever remain indebted for his tutelage. I am also very grateful for the assistance and advice provided by my doctoral committee: Dr. Krishnamurthy, Dr. Erickson, Dr. Zawodniok and Dr. Nadendla.

I have the deepest gratitude for the support and love from my family. I can never thank my parents enough for their encouragement and love. Finally, I greatly appreciate the companionship and patience of my wife, Hoda. Her support throughout my graduate studies was essential to my success. It would never have finished without her support and I will forever remain indebted to her.

TABLE OF CONTENTS

Page

SECTION

LIST OF ILLUSTRATIONS

SECTION

1. INTRODUCTION

1.1. OVERVIEW OF DYNAMICAL SYSTEMS WITH TIME DELAY

Many industrial applications, such robotic systems, suffer from internal delay due to the system components or from communication delay. This degrades the performance of the overall closed-loop system. Since delays affect the system performance, the study of the dynamic systems with time delay has received attention [1–12]. In particular, the design of optimal control schemes for such systems with input and state delays is required.

The optimal control of dynamical systems with state delay has been considered in the literature [13–18] given the system dynamics. The optimal control of linear time-invariant time-delay system with state delay using infinite horizon cost function [13] is introduced by using integro-differential or partial differential equations (PDE). To overcome the solution to PDE, the optimal control of linear time-varying delay systems over finite time horizon [19] using state feedback is dealt. However, this approach [19] still needs the solution to a delay Riccati equation and the control gain matrix which is obtained by solving ordinary differential equations (ODE). In addition, the state vector is considered measurable which can be a bottleneck in many practical applications.

By contrast, the optimal control of linear time-invariant with input delay [17] is introduced in the transformed domain [20] by converting difference-differential equations to ODE. However, optimality in the original domain is not ensured though the method did not require solution for PDEs. In [16], the robust optimal control is studied using the state transformation proposed in [20]. The optimal control in original domain is converted to the optimal control problem in the transformed domain and new cost function is derived. However, the input delay is only considered. In [18], the problem of linear quadratic suboptimal control is addressed for linear time-incvariant systems with pointwise and distributed delays. An infinite horizon performance index is defined. Although the optimal policy did not require the solution for PDEs, the solution to the ODE is still required. An iterative method is presented and an initial admissible control is required. However, in many practical applications, the full state measurement is not possible. Therefore, the design of output feedback control is of utmost importance.

The output feedback control of dynamic systems with time delay is studied in [5, 21–24]. In [21], the output feedback control of linear systems with input delay including time-varying uncertainties with disturbance is studied. A robust predictor based control is proposed which attenuates the effect of the disturbance in the presence of the system dynamics uncertainty using the output measurement, whereas the proposed approach is not optimal. The output feedback control of linear time-invariant systems with input, state and output delay is studied in [22] by using a chain of predictors under the assumption that there exists a stabilizing control for the delay-free system.

The optimal H_{∞} output feedback control of a linear time-invariant system with input delay is proposed in [25]. The optimal control is derived for non-delay case, then, the delay is added to the control input. Although the input delay is considered in the paper, the analysis of the effect for the delay and the controller design is not presented for the delayed case. Adaptive output feedback of uncertain linear time-invariant systems with input delay [5] represented as the transport PDE of delayed input is covered by using an adaptive backstepping approach whereas optimality is not studied. Conversely, the output synchronization of linear systems is proposed in [26] by using the state predictor. The objective for all the above output feedback control methods is to stabilize the time delayed linear systems but optimality is not addressed.

In practical applications, due to the uncertainties in the system dynamics, having a complete and accurate knowledge of the system dynamics is not possible. Therefore, to overcome the need for complete knowledge of the system dynamics, either learning [27] or the adaptive identifier approach [28] can be employed. Reinforcement learning (RL) based optimal adaptive control (OAC) of uncertain linear and nonlinear systems is introduced in [29]. However, state or input delays are not considered. In [30], the optimal adaptive output regulation of a linear discrete-time system with input delay is addressed using RL. An adaptive identifier is studied in [31] for linear continuous-time delay systems under the assumption that the system is asymptotically stable. The OAC of uncertain linear continuous-time systems with state and input delays is not attempted in the literature.

The event-triggered control techniques [32–36], by contrast, have been introduced to reduce the communication cost. The main aim of the event-triggered controllers is to use an aperiodic mechanism for the transmission of feedback and control policies in the communication networks without causing the overall system unstable. In the case of zero-order-hold (ZOH) based event-triggered control [32], the controller and the actuator are equipped with a ZOH device to hold the control input and the state vector, received from the controller and the sensor respectively. An event triggering condition is derived and if violated, an event is initiated. In these event sampled approaches, a controller is first designed and the threshold required for the even triggering condition is found based on the closed-loop stability.

Furthermore, many applications have nonlinear dynamics and it is essential to design the controller by considering the nonlinearities in the system dynamics. Moreover, designing the optimal control of uncertain nonlinear control systems in the presence of delays is still challenging. An optimal adaptive neural network based controller is proposed for uncertain nonlinear time-delay systems with input and state delays. The controller is designed in two conditions when the system dynamics are partially uncertain and when it is completely uncertain. A novel neural network identifier is proposed to approximate

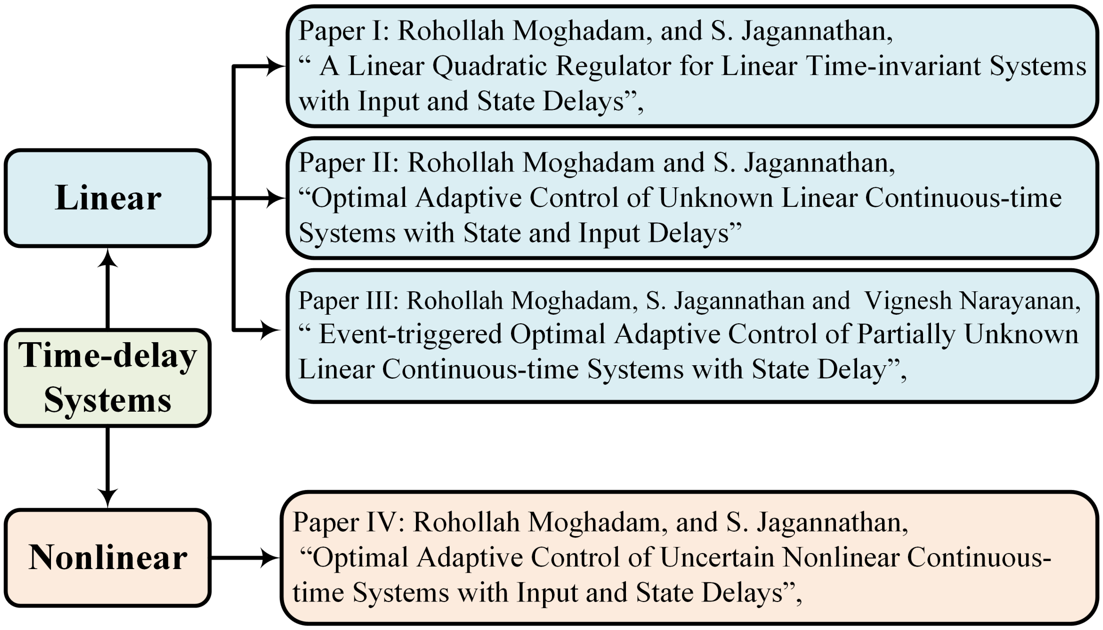


Figure 1.1. The dissertation outline.

the system dynamics. Although the optimal controller design is addressed for non-linear time-delay systems, the approach to designing the controller for time-delay systems with input and state delays using the reinforcement learning have not been investigated yet.

1.2. ORGANIZATION OF THE DISSERTATION

In this dissertation, novel optimal adaptive control of linear continuous-time delay systems with input and state delays are provided. The dissertation is presented in four papers and their relationship to one another is illustrated in Figure 1.1.

In the first paper, the optimal control of linear continuous-time systems with state and input delays by using output feedback with the objective being optimal regulation over finite and infinite horizon is presented. A quadratic cost function over finite time horizon is considered initially and optimal control policy is derived by using a time-dependent delay Riccati equation (TDDRE). For the case of optimal control over infinite horizon, a Lyapunov-Karkovskii functional is defined as a value function. Then, the Bellman type equation is formulated in terms of the value function and, using the stationarity condition, the optimal control policy is determined. Delay Algebraic Riccati Equations (DARE) is derived for output feedback and the solution is needed to determine the optimal policy.

Then, in the second paper, an optimal adaptive regulation of a completely unknown linear continuous-time system with state and input delays is addressed. First, after reviewing the OAC by using state and output feedback for the known system dynamics case, an adaptive identifier is introduced to estimate the system parameters when the dynamics are uncertain. A novel value function is then defined and, using the stationarity condition of the Bellman type equation, the optimal control input is determined. Next, the adaptive identifier and the optimal control approach are considered together to design the optimal regulator in the presence of delays. For the case of output feedback, a novel adaptive observer is proposed which not only estimates the system dynamics, but also estimates the system state. Optimal output feedback is derived through the Bellman type equation and, the boundedness of the observer, parameter estimation and output errors are presented using Lyapunov theory. The boundedness of the overall closed-loop system is demonstrated using Lyapunov stability analysis when the system matrices are uncertain under mild conditions.

The third paper presents an event-triggered optimal adaptive control (OAC) for partially unknown linear continuous time systems with state-delay by using IRL. First a quadratic cost function over an infinite time horizon is considered and a value function is defined by considering the delayed state of the system. Then the Bellman type equation is formulated and, using the stationarity condition, the optimal control is determined when the system matrices are known. A delay algebraic Riccati equation (DARE) is derived to guarantee the stability of the closed-loop system under the assumption that the system is controllable and observable. Subsequently, an actor-critic framework is utilized using the IRL approach to relax system dynamics for control policy generation. Bellman error, derived from the difference between the actual and approximated value function, is utilized to find the update law for the parameters of the critic function. Lyapunov theory is employed to demonstrate boundedness of the state vector and the parameter estimation errors.The event sampled adaptive control of time-delayed system is included. To enhance the optimality, a novel hybrid scheme is introduced, where time-varying number of iterations can be incorporated within the event sampled interval.

The fourth paper presents an optimal adaptive control (OAC) for uncertain nonlinear continuous time systems with input and state delays by using IRL. First a quadratic cost function over infinite time horizon is considered and a value function is defined by considering the delayed state of the system. Subsequently, an actor-critic framework is utilized using the IRL approach to relax the need for system dynamics for control policy generation. The critic network is approximated through a neural network and temporal difference error, derived from the difference between the actual and approximated value function, is utilized to find the update law for the weights of the critic function. Next, a novel neural network identifier is introduced to estimate the control coefficient matrix. The value function and identifier weight tuning law are derived and the boundedness of the closed-loop stability is demonstrated.

1.3. CONTRIBUTIONS

The dissertation provides contributions to the field of optimal control of time delay systems with input and state delays.

The development of time-dependent delay Riccati and delay Riccati equations for linear continuous-time systems with input and state delays and the determination of optimal control policy using output feedback are the contributions from Paper I.

In Paper II, the development of an adaptive identifier for time-delay systems, the derivation of estimated delayed algebraic Riccati equation for linear continuous-time systems with input and state delays, the design of the linear quadratic regulator in the presence of delays using state and output feedback represent the contributions of this paper.

The contributions of Paper III include the development of event-triggered based optimal adaptive control of time delay systems with state delay using integral reinforcement learning, and the study of the hybrid event-triggered approach reducing the number of transmissions of state vector to the controller while minimizing the optimal control using the game theory.

The contributions of Paper IV include the selection of a value function in terms of input and state delays, the definition of integral temporal difference error incorporating delayed state vector, the derivation of critic neural network weight tuning law using integral temporal difference error, the introduction of a neural network identifier to estimate the control coefficient matrix, and the development of finite dimensional memoryless optimal adaptive control policy for nonlinear time-delayed systems with input and state delays under partial and completely uncertain dynamics.

In all papers, the stability analysis of the closed-loop system is demonstrated in the presence of optimal control policy and verified using the simulation results.

PAPER

I. A LINEAR QUADRATIC REGULATOR FOR LINEAR TIME-INVARIANT SYSTEMS WITH INPUT AND STATE DELAY

Rohollah Moghadam and S. Jagannathan

Department of Electrical and Computer Engineering

Missouri University of Science and Technology

Rolla, Missouri 65409–0050

Email: moghadamr@umsystem.edu

ABSTRACT

In this paper, a unified linear quadratic regulator design is introduced for a linear continuous-time system with input and state delays by utilizing a quadratic cost function over an infinite horizon, using both state and output feedback. Internal and communication delays are considered as state and input delays. A Lyapunov-Karkovskii value function is defined to incorporate both state and input delays. Then, the Bellman type equation is formulated and using the stationarity condition, optimal control policy is generated. Novel delay algebraic Riccati equations are derived to find the optimal policy for both state and output feedback. Lyapunov theory is utilized to show that the optimal control renders the closed-loop system asymptotically stable. Simulation results confirm the theoretical claims.

1. INTRODUCTION

Delays are present in most industrial applications including human machine integration [1]. Since delays affect the system performance, the study of the dynamic systems with time delay has received attention [2–14]. In this regard, the design of the linear quadratic regulator for such systems in the presence of delays is recognized as of utmost importance.

The optimal control of dynamical systems with state delay has been considered in the literature [15–20] given the system dynamics. In [15], optimal control is introduced by using integro-differential or partial differential equations (PDE). To overcome the solution to PDE, the optimal control of linear time-varying delay systems over finite time horizon [21] is dealt using state feedback. However, this approach [21] still needs the solution to a delay Riccati ordinary differential equations (ODE) in order to compute control gain matrix.

By contrast, the optimal control of linear time-invariant system with input delay [19] is introduced in the transformed domain [22] by converting difference-differential equations to ODE. Though the method does not require a solution to PDEs, optimality in the original domain is not ensured. In [18], the robust optimal control is studied using the state transformation proposed in [22] for systems with input delay. In [20], the problem of linear quadratic suboptimal control is addressed for linear time-invariant systems with pointwise and distributed delays. By using an initial admissible control, an iterative optimal policy which does not require the solution of PDEs is presented, whereas, the solution for the ODE is still needed. However, in practical applications, the full state measurement is not possible and, therefore, the design of output feedback control is important.

The output feedback control of dynamic systems with time delay is studied in [4, 23–26]. In [23], robust predictor based output feedback control of linear systems with input delay including time-varying uncertainties with disturbance is studied, whereas, optimality is not addressed. The output feedback control of linear time-invariant systems with input, state and output delay is studied in [24] using a chain of predictors under the assumption that there exists a stabilizing control for the delay-free system. The optimal

H_∞ output feedback control of a linear time-invariant system with input delay is proposed in [27]. The optimal control is derived for non-delay case, then, the delay is added to the control input. Although the input delay is considered, the analyses of the effect of the delay and the controller design is not presented for the delayed case. On the other hand, the output synchronization of linear systems is proposed in [28] using the state predictor. The objective in the above output feedback control methods is to stabilize the time delayed linear systems whereas optimality is not addressed.

By contrast, this paper presents the optimal control of linear continuous-time systems with state and input delays by using both state and output feedback with the regulation as the objective. A quadratic cost function over infinite time horizon is considered. A Lyapunov-Karkovskii functional is defined as a value function. Then, the Bellman type equation (BTE) is formulated in terms of the value function and, using the stationarity condition, the optimal control policy is determined. Delay Algebraic Riccati Equations (DARE) are derived for both state and output feedback the solution of which is needed to determine the optimal policy. Finally, the stability of the closed-loop system with the optimal control policy is investigated.

The proposed approach does not require the solution to the PDE (or ODE) and any transformation and therefore it is memoryless. The contributions of the paper include the: 1) development of DARE for linear continuous-time systems with input and state delays by using Lypunov-Karkovskii value functional; 2) determination of optimal control policy using both state and output feedback; 3) stability analysis of the closed-loop system in the presence of optimal control policy; and 4) verification of the theoretical results using simulation.

The paper is organized as follows. Problem formulation is introduced in Section 2. In Section 3, the optimal control is introduced using state and output feedback. Simulation results and conclusion are provided in Sections 4 and 5, respectively.

Notations. Throughout the paper, $\|.\|$ denotes the Euclidean vector norm, I_n is the $n \times n$ identity matrix, and 0_n is the $n \times n$ zero matrix. The controllability of the pair (A, B) denotes as $ctrb(A, B)$. The rank of the matrix A is defined as $rank(A)$. Minimum and maximum singular values of the matrix A are represented by $\delta_{min}(A)$ and $\delta_{max}(A)$, respectively. The minimum and maximum eigenvalue of the matrix A is defined as $\lambda_{min}(A)$ and $\lambda_{max}(A)$, respectively. The pseudo inverse of matrix A is defined as $A^{\dagger} = (A^T A)^{-1} A$.

2. PROBLEM STATEMENT

In this section, the problem of finding the optimal control of linear continuous-time systems with input and state delays is formulated. A quadratic cost function with respect to state and control input is defined. A linear quadratic regulator is designed to drive the output vector to zero and minimizing the cost function.

Consider a linear time-invariant continuous-time system with state and input delays represented by:

$$\begin{cases} \dot{x}(t) = A_0 x(t) + A_1 x(t - d_x) + B_0 u(t) + B_1 u(t - d_u) \\ y(t) = C x(t) \end{cases}, \tag{1}$$

with

$$\begin{cases} x(\theta) = \varphi_x(\theta) \quad \theta \in [-d_x, 0] \\ u(\theta) = \varphi_u(\theta) \quad \theta \in [-d_u, 0] \end{cases} \tag{2}$$

where $x(t) \in \mathbb{R}^n$, $u(t) \in \mathbb{R}^m$ and $y(t) \in \mathbb{R}^p$ denote the state, control input and system output, with known constant state and input delays, defined as d_x and d_u, respectively. The matrices $A_0 \in \mathbb{R}^{n\times n}$, $A_1 \in \mathbb{R}^{n\times n}$, $B_0 \in \mathbb{R}^{n\times m}$, $B_1 \in \mathbb{R}^{n\times m}$ denote the drift dynamics, input matrix, respectively whereas $C \in \mathbb{R}^{p\times n}$ represents output matrix. The initial functions, φ_x and φ_u, are continuously differentiable.

Remark 1. Note that solving (1) at $t = 0$ requires the values of $x(0), x(-d_x)$ and $u(t - d_u)$. Likewise, the solution to (1) at $t = r, 0 \leqslant r < d_x, x(r), x(r - d_x)$ and $u(r - d_u)$ are required. Consequently, to find the solution to (1), one needs to have access to the state vector $x(t)$ for the time interval $-d_x \leqslant t \leqslant 0$ and the input $u(t)$ for the time interval $-d_u \leqslant t \leqslant 0$. Therefore, the differentible functions φ_x and φ_u have to be known to determine the initial conditions and their previous values.

Assumption 1. The linear time-invariant continuous-time system (1) with state and input delays is controllable and observable. In addition, the matix $C \in \mathbb{R}^{p\times n}$ satisfies $\lambda_{min}(C^T C) > 0$.

Note that, since $C^T C$ is a symmetric matrix and $\lambda_{min}(C^T C) > 0$, it becomes a symmetric positive definite matrix [29]. Therefore, pseudo inverse $C^\dagger = (C^T C)^{-1} C$ exists.

Remark 2. The linear time-invariant system (1) with input and state delays is controllable at $t_0 = 0$ if, for φ_x and $x_0 = x(0)$ and x_1, there exists $t_1 > 0$ and a control input $u(t)$ for the time interval $t \in [0, t]$, such that $x(t_1) = x_1$ [30].

Remark 3. In [3] sufficient algebraic conditions are provided as follows to check the controllability and observability of time-delay systems with input or state delays.

1. **Controllability** (Proposition 2.3 and Theorem 2.4)

   A linear continuous-time system with state delay $\dot{x}(t) = A_0 x(t) + A_1 x(t - d_x) + B_0 u(t)$ is controllable on $[0, t_1]$ for all $t_1 > n d_x$ if $rank(\tilde{Q}) = n$, where

$$\begin{aligned} &\tilde{Q} = \left[\tilde{Q}_1^1 \ldots \tilde{Q}_1^n, \tilde{Q}_2^2, \ldots, \tilde{Q}_2^n, \ldots \tilde{Q}_n^n\right] B_0 \\ &\tilde{Q}_1^1 = I, \tilde{Q}_j^{k+1} = A_0^T \tilde{Q}_j^k + A_1^T \tilde{Q}_{j-1}^k, \quad \begin{cases} j = 1, \ldots, k+1 \\ k = 1, \ldots, n-1 \end{cases} \\ &\tilde{Q}_j^k = 0, \quad j = 0 \quad or \quad j > k \end{aligned} \tag{3}$$

Therefore, if a non-delay system, i.e., $\dot{x}(t) = A_0x(t)+B_0u(t)$ is controllable on $[t_0, t_1]$, or equivalently, the pair (A_0, B_0) is controllable, then, $\dot{x}(t) = A_0x(t) + A_1x(t - d_x) + B_0u(t)$ is controllable on $[t_0, t_1]$. It has been shown in [31] (Theorems 3.6 and 3.7) that a necessary and sufficient condition for the controllability of the linear continuous time system with input delay $\dot{x}(t) = A_0x(t) + B_0u(t) + B_1u(t - d_u)$ on any $[t_0, t_1]$ with $t_1 > t_0 + d_u$ is $rank\left(ctrb\left(A_0, B_0 + e^{-A_0d_u}B_1\right)\right) = n$.

2. **Observability** (Proposition 2.4 and Theorem 2.6)

   The linear continuous-time system (1) is observable on $[0, t_1]$ for all $t_1 > nd_x$ if $rank(\bar{P}) = n$, where

$$\begin{aligned}
&\bar{P} = \left[\bar{P}_1^1 \ldots \bar{P}_1^n, \bar{P}_2^2, \ldots, \bar{P}_2^n, \ldots \bar{P}_n^n\right] C^T \\
&P_1^1 = I, \bar{P}_j^{k+1} = A_0^T\bar{P}_j^k + A_1^T\bar{P}_{j-1}^k, \quad \begin{cases} j = 1, \ldots, k+1 \\ k = 1, \ldots, n-1 \end{cases} \\
&\bar{P}_j^k = 0, \quad j = 0 \quad or \quad j > k
\end{aligned} \tag{4}$$

   Therefore, if the non-delay system of (1), i.e., $A_1 = 0$ is observable on $[t_0, t_1]$, or equivalently, the pair (A_0, C) is observable, then (1) is observable on $[t_0, t_1]$.

Remark 4. From the definition of controllability and observability in Remark 1, the time delay system (1) is observable if the pair (A_0, C) is observable. Moreover, using the definitions of controllability and as shown in [22, 32], the system (1) is controllable if the pair (A, B) is controllable, where A and B are given by $A = A_0 + e^{-Ad_x}A_1$ and $B = B_0 + e^{-Ad_u}B_1$.

Next, the problem statement is introduced. Consider the linear continuous-time system with state and input delays defined in (1). Design a control policy $u(t) = -Kx(t)$, where $K \in \mathbb{R}^{m\times n}$ is a user defined gain matrix which will be defined later, such that the closed-loop system is asymptotically stable, i.e., $\lim_{t\to\infty} x(t) \to 0$, while minimizing the quadratic cost function over infinite horizon given by

$$V(x(t)) = \int_t^\infty \left(x(t)^T \tilde{Q} x(t) + u(t)^T R u(t)\right) d\tau, \tag{5}$$

where $\tilde{Q} = C^T C \in \mathbb{R}^{n \times n}$ and $R \in \mathbb{R}^{m \times m}$ represent user defined positive definite matrices.

It has been shown in [15, 20] that the optimal control can be obtained using the following Bellman type equation (BTE) given by

$$\min_u \underbrace{\left\{ \frac{\partial V(x)}{\partial t} + x^T \tilde{Q} x + u^T R u \right\}}_{BTE} = 0, \tag{6}$$

where $\partial V(x)/\partial t$ denotes the time derivative of the value function along the system dynamics (1). Note that for the case of output feedback, the state $x(t)$ replaces the output measurement $y(t)$ in the equations (5) and (6).

In the next section, it is first assumed that the state vector of the system is measurable and, then, the optimal control policy is determined. This requirement is relaxed using output feedback in the subsequent subsection.

3. OPTIMAL REGULATION

In this section, the optimal regulator design is introduced using both state and output feedback. The state feedback case becomes a special instance of the output feedback, when C is a diagonal and full rank matrix.

3.1. STATE FEEDBACK

In this section an optimal control approach is presented for the linear continuous-time system (1) with state and input delays, using the state feedback. A novel Lyapunov-Krasovskii type functional is introduced and BTE is derived in terms of the value functional.

Then, using the BTE, the optimal control policy is found using the stationarity condition. Finally, using the Lyapunov analysis, it is shown that a closed-loop system using the optimal control policy is asymptotically stable. The following lemmas are needed to find the optimal solution to (6).

Lemma 1. [16, 17] Consider the linear continuous-time system (1) with state and input delays. Then, under mild conditions, the derivative of the delayed input, $u(t-d_u)$, is related with the non-delay input, $u(t)$, as $B_1 \partial u(t-d_u)/\partial u(t) = e^{A_0 d_u} B_1$.

Lemma 2. [29] Let M be a positive semi-definite square matrix and N be a positive definite square matrix, i.e., $v^T M v \geqslant 0$ and $v^T N v > 0$ for any vector $v \neq 0$. Then, $M + N$ is a positive definite square matrix.

To proceed, let the value function $V(x(t))$ be defined as a Lyapunov-Karkovskii functional in the form of

$$V(x) = x^T P_x x + \alpha_1 \int_{t-d_x}^{t} x^T(s) x(s) ds + \alpha_2 \int_{t-d_u}^{t} x^T(s) x(s) ds, \tag{7}$$

where $P_x \in \mathbb{R}^{n \times n}$ is a symmetric positive definite matrix and $\alpha_1 > 0$ and $\alpha_2 > 0$ are positive scalars.

Using (7), the BTE (6) can be written as

$$\begin{aligned} BTE &= (A_0 x + A_1 x(t-d_x) + B_0 u + B_1 u(t-d_u))^T P_x x \\ &+ x^T P_x (A_0 x + A_1 x(t-d_x) + B_0 u + B_1 u(t-d_u)) \\ &+ \alpha_1 x^T x - \alpha_1 x^T (t-d_x) x(t-d_x) \\ &+ \alpha_2 x^T x - \alpha_2 x^T (t-d_u) x(t-d_u) + x^T H x + u^T R u \end{aligned} . \tag{8}$$

By applying the stationarity condition [33] for the above equation, i.e., $\partial(BTE)/\partial u = 0$, we get $2Ru + 2B_0^T P_x x + 2B_1 \partial(u(t - d_u)/\partial(u(t)) = 0$. Therefore, by using the result of Lemma 1, the optimal control policy becomes

$$u^*(t) = -R^{-1} B^T P_x x(t), \tag{9}$$

where $B = B_0 + e^{A_0 d_u} B_1$, and P_x is the solution to the following delayed algebraic Riccati equation (DARE)

$$A_0^T P_x + P_x A_0 - P_x B R^{-1} B^T P_x + \tilde{Q} = 0, \tag{10}$$

with $R > 0$ and $\tilde{Q} = \alpha_1 I + \alpha_2 I + Q$, $Q \in \mathbb{R}^{n \times n}$ being a user defined positive definite matrix. By using Lemma 2, $\tilde{Q}$ becomes a positive definite matrix. The following assumption is asserted before discussing the existence and uniqueness of the solution to the DARE (10).

Assumption 2. The pair $(A_0, \sqrt{\tilde{Q}})$ is detectable and (A_0, B) is stabilizable.

Remark 5. (Existence and uniqueness of the solution to DARE). It has been shown in [34] that there exists a unique positive definite solution, i.e., $P_x > 0$, to algebraic type Ricatti equation (10), if the pair (A_0, B) is stabilizable and the pair $(A, \sqrt{\tilde{Q}})$ is detectable. Therefore, there exists a unique solution to the DARE (10) if Assumptions 1 and 2 are satisfied.

Remark 6. The design variable selection in the above theorem requires the solution to the DARE (10), which in turn needs system matrices. For instance, when $A_1 = B_1 = 0$ and the the second and third terms in the value function given in (7) are zero due to $\alpha_1 = \alpha_2 = 0$, the DARE (10) leads to the traditional algebraic Riccati equation (ARE) with $\tilde{Q} = Q = C^T C$ and $B = B_0$.

Remark 7. Note that the obtained DARE (10) depends upon the coefficient matrix of the input delay and design variable α_1, which is selected using the coefficient matrix of the state delay. Also, the control policy is memoryless and it does not need the solution to a PDE.

In the following theorem, it was shown that the control input (9) makes the continuous-time system (1) asymptotically stable in the presence of the state and input delays while minimizing the cost function (5).

Theorem 1. Consider the linear continuous-time system (1) with input and state delays. Let Assumptions 1 and 2 be satisfied and the positive definite matrix P_x be the solution to the DARE (10). Then, the control input (9) renders the closed-loop system asymptotically stable when the user defined variables are selected as $\lambda_{min}(Q) > \delta_{max}(A_1^T P_x) + \delta_{max}(P_x B R^{-1} B_1^T P_x) + \delta_{max}(P_x B R^{-1}(B - 2B_0)^T P_x)$, $\alpha_1 > \delta_{max}(A_1^T P_x)$, and $\alpha_2 > \delta_{max}(P_x B R^{-1} B_1^T P_x)$.

Proof. Consider the value function defined in (7) as a Lyapunov candidate function. Taking the derivative of (7) to get

$$\begin{aligned}\dot{V} = {} & (A_0 x + A_1 x(t-d_x) + B_0 u + B_1 u(t-d_u))^T P_x x \\ & + x^T P_x (A_0 x + A_1 x(t-d_x) + B_0 u + B_1 u(t-d_u)) \\ & + \alpha_1 x^T x - \alpha_1 x^T (t-d_x) x(t-d_x) \\ & + \alpha_2 x^T x - \alpha_2 x^T (t-d_u) x(t-d_u)\end{aligned}, \tag{11}$$

which can be written as

$$\begin{aligned}\dot{V} = {} & x^T \left(A_0^T P_x + P_x A_0 + \alpha_1 I + \alpha_2 I\right) x \\ & + 2x^T(t-d_x) A_1^T P_x x - \alpha_1 x^T(t-d_x) x(t-d_x) \\ & - \alpha_2 x^T(t-d_u) x(t-d_u) + 2u^T B_0^T P_x x \\ & + 2u^T(t-d_u) B_1^T P_x x\end{aligned}. \tag{12}$$

Now, substituting the optimal control (9) in (12), leads to

$$
\begin{aligned}
\dot{V} &= x^T \left(A_0^T P_x + P_x A_0 + \alpha_1 I + \alpha_2 I - 2 P_x B R^{-1} B_0^T P_x \right) x \\
&\quad + 2x^T(t-d_x) A_1^T P_x x - \alpha_1 x^T(t-d_x) x(t-d_x) \\
&\quad - 2x^T(t-d_u) P_x B R^{-1} B_1^T P_x x - \alpha_2 x^T(t-d_u) x(t-d_u)
\end{aligned}
\quad . \tag{13}
$$

Using the definition of B in (9), one has $B_0 = B - e^{A_0 d_u} B_1$. Then, after some manipulations, equation (13) is rewritten as

$$
\begin{aligned}
&\dot{V} = \\
&x^T \left(A_0^T P_x + P_x A_0 + \alpha_1 I + \alpha_2 I - P_x B R^{-1} B^T P_x \right) x \\
&\quad + x^T P_x B R^{-1} (B - 2B_0)^T P_x x \\
&\quad + 2x^T(t-d_x) A_1^T P_x x - \alpha_1 x^T(t-d_x) x(t-d_x) \\
&\quad - 2x^T(t-d_u) P_x B R^{-1} B_1^T P_x x - \alpha_2 x^T(t-d_u) x(t-d_u)
\end{aligned}
\quad . \tag{14}
$$

Next, using the DARE (10), (14) becomes

$$
\begin{aligned}
\dot{V} &= -x^T Q x + x^T P_x B R^{-1} (B - 2B_0)^T P_x x \\
&\quad + 2x^T(t-d_x) A_1^T P_x x - \alpha_1 x^T(t-d_x) x(t-d_x) \\
&\quad - 2x^T(t-d_u) P_x B R^{-1} B_1^T P_x x - \alpha_2 x^T(t-d_u) x(t-d_u)
\end{aligned}
\quad . \tag{15}
$$

Using the Young's inequality for the cross product terms, (15) to get

$$
\begin{aligned}
\dot{V} &\leqslant -\lambda_{\min}(Q) \|x\|^2 + x^T P_x B R^{-1} (B - 2B_0)^T P_x x \\
&\quad + \delta_{max}(A_1^T P_x) \|x\|^2 + \delta_{max}(P_x B R^{-1} B_1^T P_x) \|x\|^2 \\
&\quad + \delta_{max}(A_1^T P_x) \|x(t-d_x)\|^2 - \alpha_1 \|x(t-d_x)\|^2 \\
&\quad + \delta_{max}(P_x B R^{-1} B_1^T P_x) \|x(t-d_u)\|^2 - \alpha_2 \|x(t-d_u)\|^2
\end{aligned}
\quad . \tag{16}
$$

The above equation can be expressed as

$$
\begin{aligned}
\dot{V} \leqslant & -\left(\lambda_{\min}(Q) - \delta_{max}\left(A_1^T P_x\right) - \delta_{max}\left(P_x B R^{-1} B_1^T P_x\right)\right. \\
& \left. -\delta_{max}\left(P_x B R^{-1} (B - 2B_0)^T P_x\right)\right) \|x\|^2 \\
& -\left(\alpha_1 - \delta_{max}(A_1^T P_x)\right) \|x(t-d_x)\|^2 \\
& -\left(\alpha_2 - \delta_{max}\left(P_x B R^{-1} B_1^T P_x\right)\right) \|x(t-d_u)\|^2
\end{aligned} . \tag{17}
$$

Therefore, one can conclude that if conditions in the statement of the theorem are met, $\dot{V} < 0$. This states that the control input (9) makes the continuous-time system (1) with input and state delays asymptotically stable. ■

Remark 8. Note that the conditions in the statement of Theorem 2 depend on the design parameters Q, α_1 and α_2. Simply, the designer can choose the design variables to ensure the boundary conditions are satisfied.

Remark 9. Alternatively, for example in [18, 19], the optimal control is obtained in the transformed domain. The state transformation method presented in [22] is revisited and it is given by

$$
\begin{aligned}
z(t) = x(t) & + \int_{t-d_x}^{t} e^{-A(s+d_x-t)} A_1 x(s) ds \\
& + \int_{t-d_u}^{t} e^{-A(s+d_u-t)} B_1 u(s) ds
\end{aligned} .
$$

Then, the dynamics of the original continuous-time system (1) with input and state delays in the transformed domain becomes $\dot{z}(t) = Az(t) + Bu(t)$, where $A \in \mathbb{R}^{n\times n}$ is given by $A = A_0 + e^{-Ad_x} A_1$ and $B = B_0 + e^{-Ad_u} B_1$. It can be seen that the state transformation converts the infinite dimensional delay differential equations to a finite ordinary differential equation. In [19], the state transformation is used to find the optimal control of a linear time delay system with input delay. However, finding the optimality in the transformed domain does not guarantee the optimality in the original domain. In [18], the optimal tracking error using the state transformation is considered. Although the relation between the optimality in the transformed domain and the original domain is discussed, the proposed approach is only applicable for the case of input delay.

It has been shown in Theorem 1 that the optimal regulation using state feedback, derived from the stationarity condition, makes the closed-loop system (1) asymptotically stable, if Assumptions 1 and 2 are held and the DARE (10) is satisfied. In the next subsection, the measured output vector is used to obtain optimal regulation.

3.2. OUTPUT FEEDBACK

An optimal output feedback approach is proposed for a linear continuous-time system (1) with state and input delays. The value function is defined as a function of the system output. Then, the BTE is found using the corresponding value function. The optimal control is determined using the stationarity condition which results in the derivation of an output DARE. Finally, the asymptotic convergence of the system output to zero is shown, using Lyapunov theory.

Define the quadratic cost function over infinite horizon as

$$V(y(t)) = \int_t^{\infty} \left(y(t)^T \bar{Q} y(t) + u(t)^T R u(t)\right) d\tau \,,$$

and the corresponding BTE for the case of output feedback is considered as

$$\min_u \left\{ \frac{\partial V(y)}{\partial t} + y^T \bar{Q} y + u^T R u \right\} = 0.$$

The value function $V(y(t))$ for the case of output feedback is defined as

$$V(y) = y^T P y + \alpha_1 \int_{t-d_x}^{t} y^T(s) y(s) ds + \alpha_2 \int_{t-d_u}^{t} y^T(s) y(s) ds, \tag{18}$$

where $P \in \mathbb{R}^{n\times n}$ is a positive definite matrix and the scalars $\alpha_1 > 0$ and $\alpha_2 > 0$. Now, using (18), the BTE (6) can be written as

$$\begin{aligned} BTE &= (A_0x + A_1x(t-d_x) + B_0u + B_1u(t-d_u))^T C^T P y \\ &+ y^T PC(A_0x + A_1x(t-d_x) + B_0u + B_1u(t-d_u)) \\ &+ \alpha_1 y^T y - \alpha_1 y^T(t-d_x)y(t-d_x) \\ &+ \alpha_2 y^T y - \alpha_2 y^T(t-d_u)y(t-d_u) + y^T\bar{Q}y + u^T R u \end{aligned} . \tag{19}$$

By applying the stationarity condition $\partial(BTE)/\partial u = 0$ on (19) and using Lemma 1, one has

$$u^*(t) = -R^{-1}B^T C^T P y(t), \tag{20}$$

where $B = B_0 + e^{A_0 d_u}B_1$, $P \in \mathbb{R}^{n\times n}$ is obtained as $P_y = C^T P C$ with $P_y \in \mathbb{R}^{n\times n}$ being the solution to the following output DARE as given by

$$A_0^T P_y + P_y A_0 - P_y B R^{-1} B^T P_y + \bar{Q} = 0, \tag{21}$$

with $R > 0$, $\bar{Q} = (\alpha_1 I + \alpha_2 I)C^T C + Q$, $Q \in \mathbb{R}^{n\times n}$ a user defined positive definite matrix. Since C is a full rank design matrix, then $C^T C$ becomes positive semi-definite as for any vector v, one has $v^T C^T C v = (Cv)^T(Cv) \geqslant 0$. Moreover, $\alpha_1 > 0$, $\alpha_2 > 0$, and Q is a positive definite matrix and based on Lemma 2, $\bar{Q}$ becomes a positive definite matrix.

Remark 10. (Existence and uniqueness of the solution to the output DARE). To guarantee the existence and uniqueness positive definite solution to (21), i.e., $P_y > 0$, the pairs (A_0, B) and $(A_0, \sqrt{\bar{Q}})$ must be stabilizable and detectable, respectively. Similar to Remark 5, this means a unique solution exists to the output DARE (21) if Assumptions 1 and 2 are fulfilled.

Remark 11. Note that it has been shown that the solution to (21) is a symmetric positive definite matrix, i.e. $P_y > 0$. Now, it is shown that matrix P is PD. To this end, from the definition of P one has $P = C^{\dagger}P_y(C^{\dagger})^T$. Using the definition of positive semi-definiteness, let z be a vector, then, $z^T C^{\dagger} P_y (C^{\dagger})^T z \geqslant 0$ can be written as $\bar{z}^T P \bar{z}$ where $\bar{z} = zC^{\dagger}$. Since $P_y > 0$, then, $\bar{z}^T P \bar{z} \geqslant 0$. Now, assume $\bar{z}^T P \bar{z} = 0$ and since $P > 0$, it implies that $\bar{z} = 0$, and consequently, $C^{\dagger} z = 0$. On the other hand, $C^{\dagger}$ is a full rank matrix which implies $ker(C^{\dagger}) = 0$. Finally, $C^{\dagger} z = 0$ concludes $z = 0$. This means $P > 0$ is a positive definite.

Theorem 2 shows that the control input (20) makes the closed-loop linear continuous time system asymptotically stable in the presence of the state and input delays.

Theorem 2. Consider the linear continuous-time system (1) with input and state delays a. Let Assumption 1 through 2 be satisfied and the positive definite matrix P_y be the solution to the output DARE (21). Then, the control input (20) renders the closed-loop system asymptotically stable if the user defined variables are selected such that $\alpha_2 > \frac{\delta_{max}(P_y B R^{-1} B_1^T P_y)}{\lambda_{min}(C^T C)}$, $\alpha_1 > \frac{\delta_{max}(A_1^T P_y)}{\lambda_{min}(C^T C)}$, and $\lambda_{min}(Q) > \delta_{max}\left(P_y B R^{-1}(B - 2B_0)^T P_y\right) + \delta_{max}(A_1^T P_y) + \delta_{max}(P_y B R^{-1} B_1^T P_y)$.

Proof. Let the Lyapunov candidate function be defined as (18). Taking the derivative of (18) to obtain

$$\begin{aligned}
\dot{V} = {} & (A_0 x + A_1 x(t-d_x) + B_0 u + B_1 u(t-d_u))^T C^T P y \\
& + y^T P C (A_0 x + A_1 x(t-d_x) + B_0 u + B_1 u(t-d_u)) \\
& + \alpha_1 y^T y - \alpha_1 y^T(t-d_x) y(t-d_x) \\
& + \alpha_2 y^T y - \alpha_2 y^T(t-d_u) y(t-d_u)
\end{aligned} \quad . \tag{22}$$

Using the definition $P_y = C^T P C$, equation (22) can be written as

$$\begin{aligned}\dot{V} &= x^T \left(A_0^T P_y + P_y A_0 + (\alpha_1 + \alpha_2) C^T C \right) x \\ &+ 2x^T(t - d_x) A_1^T P_y x - \alpha_1 x^T(t - d_x) C^T C x(t - d_x) \\ &- \alpha_2 x^T(t - d_u) C^T C x(t - d_u) + 2u^T B_0^T P_y x \\ &+ 2u^T(t - d_u) B_1^T P_y x\end{aligned} . \tag{23}$$

Now, substituting the optimal control (20) in (23) and using the definition of $B_0 = B - e^{A_0 d_u} B_1$, one has

$$\begin{aligned}\dot{V} &= x^T \left(A_0^T P_y + P_y A_0 + (\alpha_1 + \alpha_2) C^T C \right. \\ &\left. -P_y B R^{-1} B^T P_y \right) x + x^T P_y B R^{-1} (B - 2B_0)^T P_y x \\ &+ 2x^T(t - d_x) A_1^T P_y x - \alpha_1 x^T(t - d_x) C^T C x(t - d_x) \\ &- 2x^T(t - d_u) P_y B R^{-1} B_1^T P_y x - \alpha_2 x^T(t - d_u) C^T C x(t - d_u)\end{aligned} . \tag{24}$$

Using the output DARE (21), equation (24) becomes

$$\begin{aligned}\dot{V} &= -x^T Q x + x^T P_y B R^{-1} (B - 2B_0)^T P_y x \\ &+ 2x^T(t - d_x) A_1^T P_y x - \alpha_1 x^T(t - d_x) C^T C x(t - d_x) \\ &- 2x^T(t - d_u) P_y B R^{-1} B_1^T P_y x - \alpha_2 x^T(t - d_u) C^T C x(t - d_u)\end{aligned} . \tag{25}$$

Next, utilizing the Young's inequality, equation (25) results in the first derivative of the Lypunov function candidate being

$$\begin{aligned}\dot{V} &\leqslant -\lambda_{\min}(Q) \|x\|^2 + x^T P_y B R^{-1} (B - 2B_0)^T P_y x \\ &+ \delta_{max}(A_1^T P_y) \|x\|^2 + \delta_{max}(P_y B R^{-1} B_1^T P_y) \|x\|^2 \\ &+ \delta_{max}(A_1^T P_x) \|x(t - d_x)\|^2 - \alpha_1 \lambda_{\min}(C^T C) \|x(t - d_x)\|^2 \\ &+ \delta_{max}(P_y B R^{-1} B_1^T P_y) \|x(t - d_u)\|^2 - \alpha_2 \lambda_{\min}(C^T C) \|x(t - d_u)\|^2\end{aligned} , \tag{26}$$

which can be expressed as

$$
\begin{aligned}
\dot{V} \leqslant & -\left(\lambda_{\min}(Q)-\delta_{max}\left(A_1^T P_y\right)-\delta_{max}\left(P_y B R^{-1} B_1^T P_y\right)\right. \\
& \left.-\delta_{max}\left(P_y B R^{-1}(B-2B_0)^T P_y\right)\right)\left\|C^{\dagger} y\right\|^2 \\
& -\left(\alpha_1-\frac{\delta_{max}(A_1^T P_y)}{\lambda_{\min}(C^T C)}\right)\left\|C^{\dagger} y(t-d_x)\right\|^2 \\
& -\left(\alpha_2-\frac{\delta_{max}\left(P_y B R^{-1} B_1^T P_y\right)}{\lambda_{\min}(C^T C)}\right)\left\|C^{\dagger} y(t-d_u)\right\|^2
\end{aligned} \quad . \tag{27}
$$

Therefore, one can conclude that if conditions in the statement of the theorem are met, $\dot{V} < 0$. This states that the optimal control input (20), eventually guarantees the asymptotic stability of the closed-loop system (1) with input and state delays. ■

In the next section, simulation results are provided to confirm the effectiveness of the proposed approach.

4. SIMULATION RESULTS

In this section, two examples are provided one using state and another via output feedback. For the case of state feedback, a chemical reactor is considered as an example and the state vector is considered measurable for different values of delays. Subsequently, an example is provided for the case of the output feedback.

4.1. EXAMPLE 1. STATE FEEDBACK APPROACH

In this section, the presented control approach is applied to the application of a chemical reactor to reveal its validity. The chemical refining process presented in [35], which is shown in Figure 1, with transport and input delays is considered. The process includes the production of P due to the chemical reaction of the raw materials A and B.

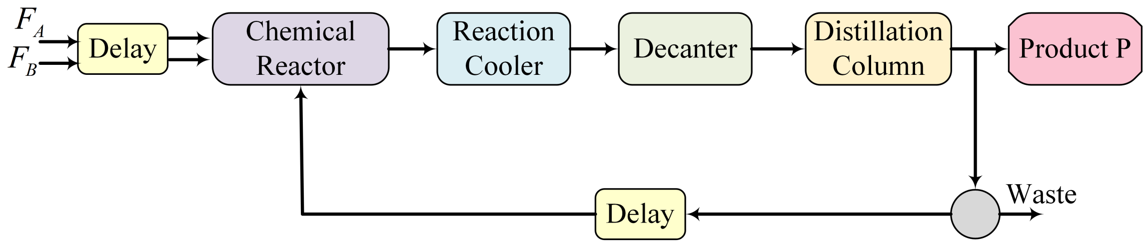


Figure 1. The chemical reactor refining plant diagram.

The linearized representation of a chemical reactor is given by

$$\begin{cases} \dfrac{da}{dt} = -4.93a(t) + 1.92a(t-1) - 1.01b(t) + u_1(t-1) \\ \dfrac{db}{dt} = -3.20a(t) - 5.30b(t) + 1.92b(t-1) - 12.8c(t) + u_2(t-1) \\ \dfrac{dc}{dt} = 6.40a(t) + 0.347b(t) - 32.5c(t) + 1.87c(t-1) - 1.04p(t) \\ \dfrac{dp}{dt} = 0.833b(t) + 11.0c(t) - 3.96p(t) + 0.724p(t-1) \end{cases}, \tag{28}$$

where $a(t)$ and $b(t)$ denote the deviations in the weight of reactants A and B, respectively. The state $c(t)$ and $p(t)$ represents the deviation in the weight of an intermediate product C and P respectively. The control inputs to the chemical reactor are given by $u_1(t) = \frac{\delta F_A}{6V_R}$ and $u_2(t) = \frac{\delta F_B}{6V_R}$ where δF_A and δF_B are the deviations of the feed rates of the materials A and B from their nominal values, respectively and V_R is the pound-volume of the chemical reactor. Note that in the general formulation of the chemical reactor presented in [35], there is no delay in the control input and the approach is limited to the transport delay in the state. However, in this paper, to show the validity of the proposed approach, the delay is also assumed to be in the feed rate of the chemical reactor which is considered as an input delay. The proposed representation of a chemical reactor with state and input delays is a more generic and practical representation than delay in the state alone.

Now, using the state representation of (28), one has

$$\dot{x}(t) = A_0x(t) + A_1x(t-1) + B_1u(t-1), \tag{29}$$

where

$$A_0 = \begin{bmatrix} -4.93 & -1.01 & 0 & 0 \\ -3.20 & -5.30 & -12.8 & 0 \\ 6.40 & 0.347 & -32.5 & -1.04 \\ 0 & 0.833 & 11.0 & -3.96 \end{bmatrix},$$

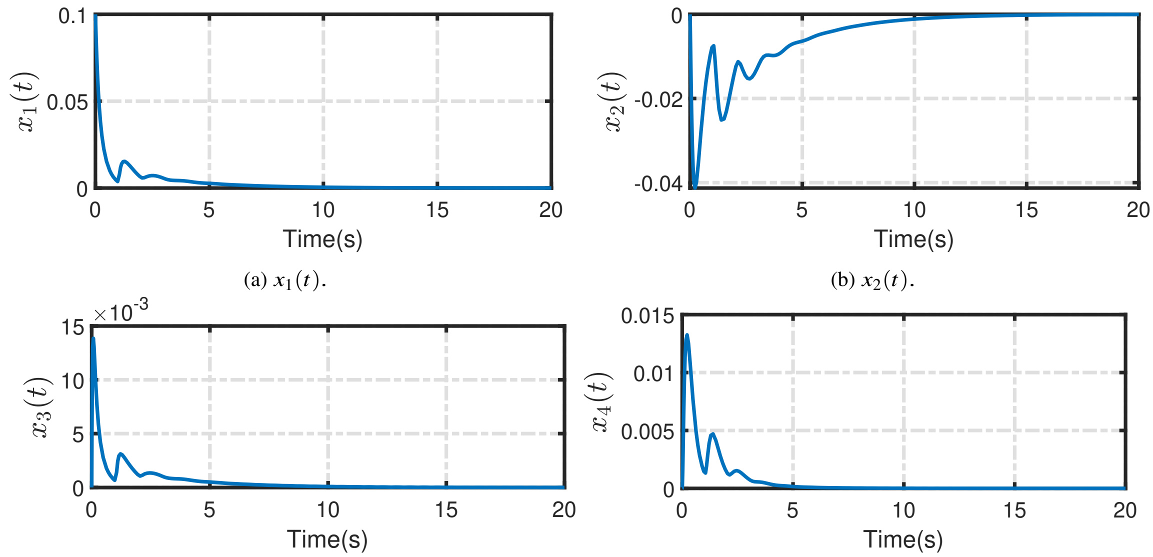
$$A_1 = \begin{bmatrix} 1.92 & 0 & 0 & 0 \\ 0 & 1.92 & 0 & 0 \\ 0 & 0 & 1.87 & 0 \\ 0 & 0 & 0 & 0.724 \end{bmatrix}, \quad B_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix},$$

$x = [x_1, x_2, x_3, x_4]^T$ being the state vector of the reactor, with $x_1(t) = a(t)$, $x_2(t) = b(t)$, $x_3(t) = c(t)$ and $x_4(t) = p(t)$. From (29), one can see that since the system dynamics do not depend upon the non-delay input $u(t)$, the input matrix $B_0 = 0$, and therefore, $B = e^{A_0 d_u} B_1$.

The simulation parameters are chosen as

$$\tilde{Q} = \begin{bmatrix} 10 & 0 & 0 & 0 \\ 0 & 10 & 0 & 0 \\ 0 & 0 & 10 & 0 \\ 0 & 0 & 0 & 10 \end{bmatrix}, \quad R = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Then, the conditions presented in Theorem 1 become $\alpha_1 > \lambda_{max}(A_1^T P) = 4.73$, $\alpha_2 > \delta_{max}(P_x B R^{-1} B_1^T P_x) = 0.02$ and $\lambda_{min}(Q) > \delta_{max}(A_1^T P_x) + \delta_{max}(P_x B R^{-1} B_1^T P_x) + \delta_{max}\left(P_x B R^{-1}(B - 2B_0)^T P_x\right) = 4.76$. The above conditions imply that $\lambda_{min}(\tilde{Q}) > 9.5$ which is $\lambda_{min}(\tilde{Q}) = 10$ and satisfies the conditions of Theorem 1. The solution to (10)



(a) $x_1(t)$.

(b) $x_2(t)$.

(c) $x_3(t)$.

(d) $x_4(t)$.

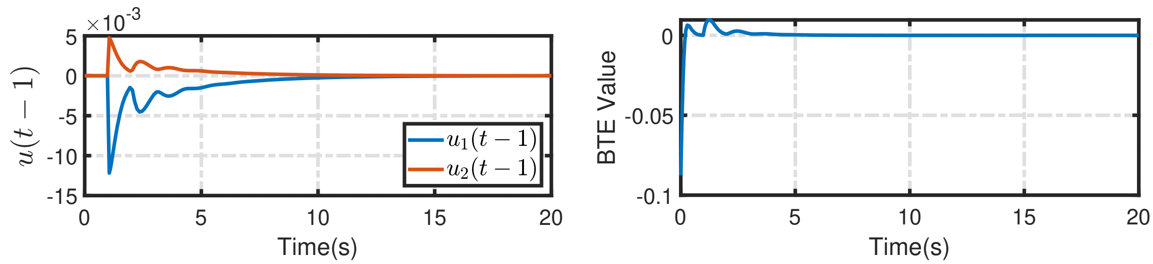
Figure 2. Performance of the controller when input delay $d_u = 1$ and state delay $d_x = 1$.

becomes

$$P = \begin{bmatrix} 1.9588 & -0.71189 & 0.37166 & 0.11952 \\ -0.71189 & 1.08 & -0.32452 & 0.14111 \\ 0.37166 & -0.32452 & 0.38217 & 0.29697 \\ 0.11952 & 0.14111 & 0.29697 & 1.1846 \end{bmatrix}.$$

The eigenvalues of the P are given by $0.15, 0.69, 1.29, 2.47$ which are positive and confirm the positive definiteness of P matrix. The performance of the proposed control input (9) for the chemical reactor (29) can be seen in Figure 2. One can observe that the state vector converges to zero asymptotically which confirms the results of Theorem 2.

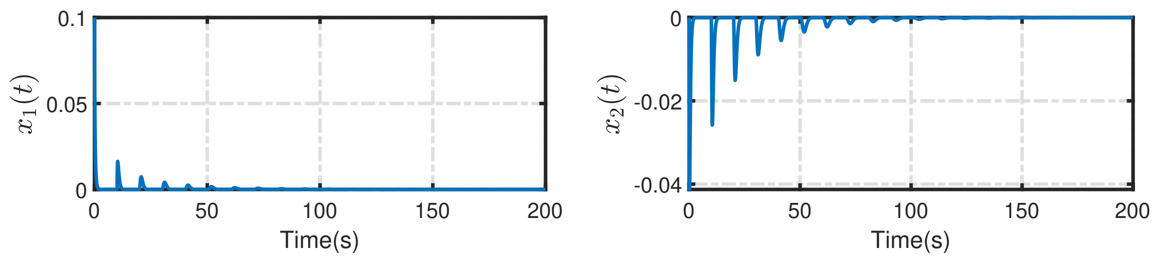
Using the proposed approach there is no need to solve PDEs, which is computationally expensive to find the optimal solution for time delay systems with state and input delays. In Figure4 the value of the the input delay and the state delay are considered as $d_u = 5$ and $d_x = 10$, respectively. The control input and the BTE value are also shown in Figure 5.



(a) Control input.

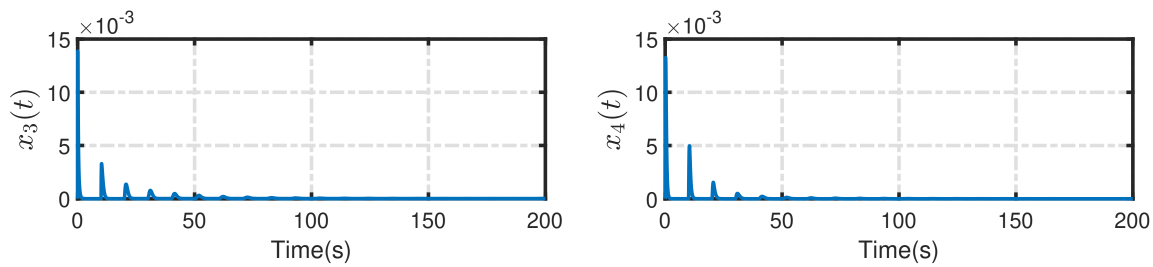
(b) Bellman type equation value.

Figure 3. Control input and BTE trajectories when input delay $d_u = 1$ and state delay $d_x = 1$.



(a) $x_1(t)$.

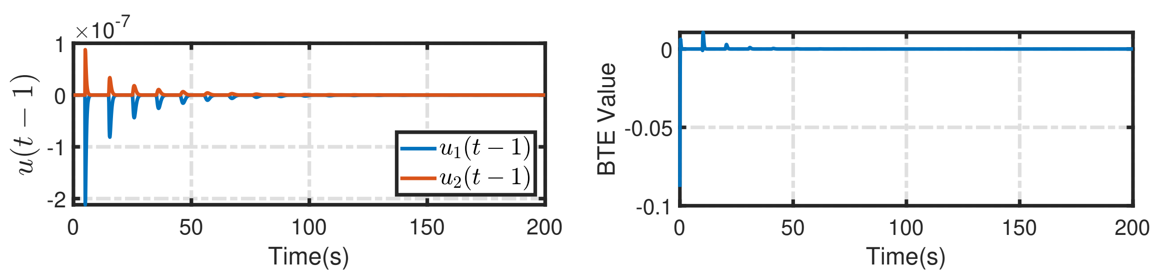
(b) $x_2(t)$.



(c) $x_3(t)$.

(d) $x_4(t)$.

Figure 4. State trajectory when input delay $d_u = 5$ and state delay $d_x = 10$.



(a) Control input.

(b) Bellman type equation value.

Figure 5. The control input and BTE trajectories when the input delay $d_u = 5$ and the state delay $d_x = 10$.

4.2. EXAMPLE 2. OUTPUT FEEDBACK APPROACH

Consider the linear continuous-time system with state and input delays given by

$$\begin{cases} \dot{x}(t) = A_0x(t) + A_1x(t-1) + B_0u(t) + B_1u(t-1) \\ y(t) = Cx(t) \end{cases}, \tag{30}$$

where

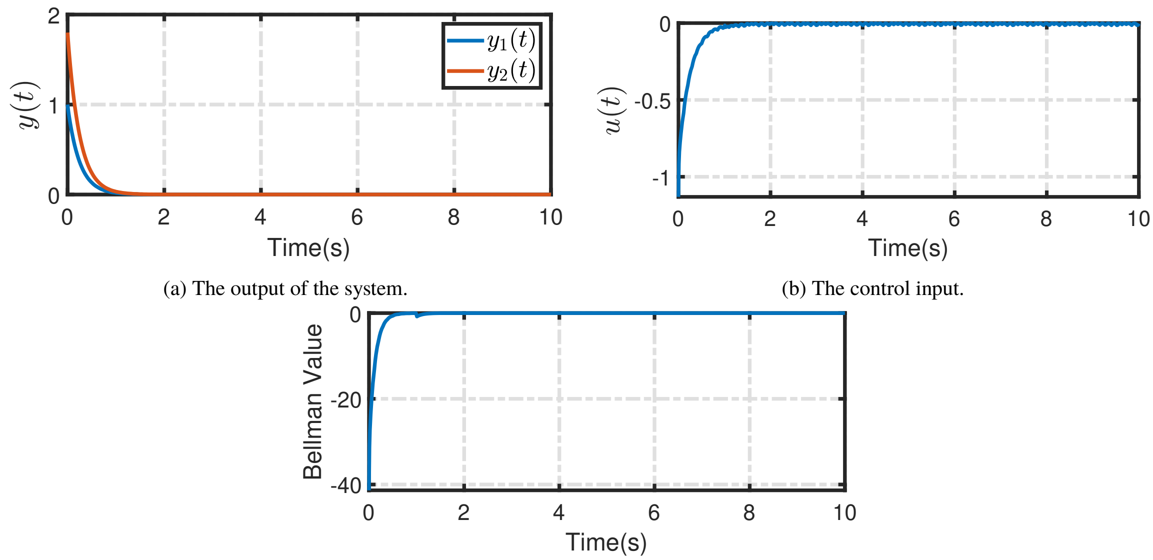
$$A_0 = \begin{bmatrix} 1 & 4 \\ 1 & -5 \end{bmatrix}, A_1 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

$$B_0 = \begin{bmatrix} 10 \\ 0 \end{bmatrix}, B_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, C = \begin{bmatrix} 1 & 0 \\ 0.8 & 1 \end{bmatrix}.$$

The history profiles for state and input are considered as $\varphi_x(\theta) = [1, 1]^T, \theta \in [-1, 0]$ and $\varphi_u(\theta) = 0, \theta \in [-1, 0]$. Next, in order to determine the design parameters as needed for the optimal adaptive regulator, the bounds in Assumption 2 (or nominal model of known dynamics) are considered. In addition, the following design parameters are selected: $Q = 10I_2$ and $R = 1, \alpha_1 = 1$ and $\alpha_2 = 3$ which results in $\tilde{Q} = 4(C^TC) + Q = \begin{bmatrix} 15.4 & 2.7 \\ 2.7 & 13.3 \end{bmatrix}$. Then, the solution to the DARE (21) becomes

$$P^* = \begin{bmatrix} 0.28 & 0.03 \\ 0.03 & 1.25 \end{bmatrix}$$

One can compute $\alpha_1 = 1 > (\delta_{max}(A_1P)/\lambda_{min}(C^TC)) = 0.62$ and $\alpha_2 = 3 > (\delta_{max}(PB_yR_{-1}B_1^TP)/\lambda_{min}(C^TC)) = 2.7$ and $\lambda_{min}(Q) = 10 > \delta_{max}\left(PBR^{-1}(B - 2B_0)^TP\right) + \delta_{max}(A_1^TP) + \delta_{max}(PBR^{-1}B_1^TP) = 8.8$. The design parameters satisfy the bound conditions provided in Theorem 2.



(a) The output of the system.

(b) The control input.

(c) The Bellman type equation value.

Figure 6. Performance of the system with the output feedback controller.

The performance of the output feedback controller (20) is shown in Figure 6. From Figure 6a wherein the output vector converges to zero within 2 seconds. The optimal control input and the value of the Bellman type equation (19) is displayed in Figures 6b and 6c, respectively. These results substantiate the theoretical claim in Theorem 2.

5. CONCLUSIONS

In this paper, by utilizing a quadratic cost function over infinite horizon the optimal regulation of linear continuous-time systems with state and input delays is addressed. The Lyapunnov -Karkovskii value function and Bellman type equation facilitated the optimal control policy through delay algebraic Ricatti equations without using state transformation or needing to solve PDE or ODE. The optimality is shown in the original domain using state and output feedback provided the system matrices are known. Lyapunov analysis demonstrated the asymptotic stability of the closed-loop system by using the optimal control input provided the design variables are selected per the theoretical analysis and the system under consideration is controllable and observable. Simulation examples help demonstrate the validity of the proposed approach.

REFERENCES

[1] D. McRuer. Human dynamics in man-machine systems. Automatica, 16(3):237 – 253, 1980.

[2] Keqin Gu, Jie Chen, and Vladimir L Kharitonov. Stability of time-delay systems. Springer Science & Business Media, 2003.

[3] Emilia Fridman. Introduction to Time-delay Systems: Analysis and Control. Springer, 2014.

[4] Y. Zhu, M. Krstic, and H. Su. Adaptive output feedback control for uncertain linear time-delay systems. IEEE Transactions on Automatic Control, 62(2):545–560, Feb 2017.

[5] A. Ponomarev. Reduction-based robustness analysis of linear predictor feedback for distributed input delays. IEEE Transactions on Automatic Control, 61(2):468–472, Feb 2016.

[6] Z. Artstein. Linear systems with delayed controls: A reduction. IEEE Transactions on Automatic Control, 27(4):869–879, August 1982.

[7] X. Cai, N. Bekiaris-Liberis, and M. Krstic. Input-to-state stability and inverse optimality of linear time-varying-delay predictor feedbacks. IEEE Transactions on Automatic Control, 63(1):233–240, Jan 2018.

[8] V. L. Kharitonov, S. . Niculescu, J. Moreno, and W. Michiels. Static output feedback stabilization: necessary conditions for multiple delay controllers. IEEE Transactions on Automatic Control, 50(1):82–86, Jan 2005.

[9] D. Bresch-Pietri and M. Krstic. Delay-adaptive predictor feedback for systems with unknown long actuator delay. IEEE Transactions on Automatic Control, 55(9):2106–2112, Sep. 2010.

[10] S. Su, Y. Wei, and Z. Lin. Stabilization of discrete-time linear systems with an unknown time-varying delay by switched low-gain feedback. IEEE Transactions on Automatic Control, 64(5):2069–2076, May 2019.

[11] Kun Liu, Anton Selivanov, and Emilia Fridman. Survey on time-delay approach to networked control. Annual Reviews in Control, 48:57 – 79, 2019.

[12] Emilia Fridman. A refined input delay approach to sampled-data control. Automatica, 46(2):421 – 427, 2010.

[13] Kun Liu, Vladimir Suplin, and Emilia Fridman. Stability of linear systems with general sawtooth delay. IMA Journal of Mathematical Control and Information, 27(4):419–436, 2010.

[14] A. T. Koru, T. Yucelen, R. Sipahi, A. Ramírez, and K. M. Dogan. Stability of human-in-the-loop multiagent systems with time delays. In 2019 American Control Conference (ACC), pages 4854–4859, 2019.

[15] D. Ross and I. Flügge-Lotz. An optimal control problem for systems with differential-difference equation dynamics. SIAM Journal on Control, 7(4):609–623, 1969.

[16] Michael Basin, Jesus Rodriguez-Gonzalez, and Leonid Fridman. Optimal and robust control for linear state-delay systems. Journal of the Franklin Institute, 344(6):830 – 845, 2007.

[17] M. Basin and J. Rodriguez-Gonzalez. Optimal control for linear systems with multiple time delays in control input. IEEE Transactions on Automatic Control, 51(1):91–97, Jan 2006.

[18] Yusheng Zhou and Zaihua Wang. A robust optimal trajectory tracking control for systems with an input delay. Journal of the Franklin Institute, 353(12):2627 – 2649, 2016.

[19] Guo-Ping Cai, Jin-Zhi Huang, and Simon X Yang. An optimal control method for linear systems with time delay. Computers & Structures, 81(15):1539 – 1546, 2003.

[20] O. Santos, S. Mondié, and V.L. Kharitonov. Linear quadratic suboptimal control for time delays systems. International Journal of Control, 82(1):147–154, 2009.

[21] Michael Basin and Jesus Rodriguez-Gonzalez. A closed-form optimal control for linear systems with equal state and input delays. Automatica, 41(5):915 – 920, 2005.

[22] Y. Fiagbedzi and A. Pearson. Feedback stabilization of linear autonomous time lag systems. IEEE Transactions on Automatic Control, 31(9):847–855, Sep. 1986.

[23] Shoulin Hao, Tao Liu, and Bin Zhou. Output feedback anti-disturbance control of input-delayed systems with time-varying uncertainties. Automatica, 104:8 – 16, 2019.

[24] Filippo Cacace and Alfredo Germani. Output feedback control of linear systems with input, state and output delays by chains of predictors. Automatica, 85:455 – 461, 2017.

[25] Le Van Hien and Mai Thi Hong. An optimization approach to static output-feedback control of lti positive systems with delayed measurements. Journal of the Franklin Institute, 2019.

[26] Saeed Ahmed, Frédéric Mazenc, and Hitay Özbay. Dynamic output feedback stabilization of switched linear systems with delay via a trajectory based approach. Automatica, 93:92 – 97, 2018.

[27] Chi-Chang Lin, Chang-Ching Chang, and Huang-Lin Chen. Optimal h_∞ output feedback control systems with time delay. Journal of Engineering Mechanics, 132(10):1096–1105, 2006.

[28] Se Young Yoon and Zongli Lin. Robust output regulation of linear time-delay systems: A state predictor approach. International Journal of Robust and Nonlinear Control, 26(8):1686–1704, 2016.

[29] Gilbert Strang. Introduction to Linear Algebra. Cambridge Press, 2016.

[30] Olivier Sename, Jean-François Lafay, and Rabah Rabah. Controllability indices of linear systems with delays. Kybernetika, 31(6):559–580, 1995.

[31] H. T. Banks, M. Q. Jacobs, and M. R. Latina. The synthesis of optimal controls for linear, time-optimal problems with retarded controls. Journal of Optimization Theory and Applications, 8(5):319–366, Nov 1971.

[32] W. Kwon and A. Pearson. Feedback stabilization of linear systems with delayed control. IEEE Transactions on Automatic Control, 25(2):266–269, April 1980.

[33] Frank L Lewis, Draguna Vrabie, and Vassilis L Syrmos. Optimal Control. John Wiley & Sons, 2012.

[34] RE KALMAN. Contribution to the theory of optimal control. Bol. Soc. Mat. Mexicana, 5:102–119, 1960.

[35] D. Ross. Controller design for time lag systems via a quadratic criterion. IEEE Transactions on Automatic Control, 16(6):664–672, December 1971.

II. OPTIMAL ADAPTIVE OUTPUT FEEDBACK CONTROL OF UNCERTAIN LINEAR CONTINUOUS-TIME SYSTEMS WITH STATE AND INPUT DELAYS

Rohollah Moghadam and S. Jagannathan

Department of Electrical and Computer Engineering

Missouri University of Science and Technology

Rolla, Missouri 65409–0050

Email: moghadamr@umsystem.edu

ABSTRACT

In this paper, near optimal regulation of linear continuous systems with uncertain dynamics in the presence of input and state delays using output feedback is addressed. First, a novel Lyapunov-Karakovskii value functional is defined and a Bellman type equation (BTE) is derived for the case of known dynamics. By applying the stationarity condition on the BTE, the optimal control policy is derived in terms of the solution to a delayed algebraic Ricatti equation (DARE) and measured output. Next for the case of unknown dynamics and to obtain estimated optimal control policy, a novel adaptive observer is proposed to estimate both the system dynamics and state vector in the presence of state and input delays. The unknown parameters of the adaptive observer are tuned using measured output. Subsequently, near optimal adaptive regulator using adaptive observer and measured output is derived using estimated DARE. The performance of the adaptive observer and regulator are demonstrated using Lyapunov analysis where it is shown that the estimated optimal control attains true optimal control policy asymptotically under a persistency of excitation condition for the ideal case and bounded for the general case. Simulation example is provided to verify the effectiveness of the proposed approach.

1. INTRODUCTION

Delays are found in the state and input vectors of many industrial applications and as a result, the performance of the overall closed-loop system deteriorates. Therefore, the control of time-delay linear systems has been investigated in the literature [1–5] when the dynamics are known. In particular, the optimal control of dynamical systems with state delay is studied in [6–8]. Since partial differential equations (PDEs) are difficult to handle in the controller design [6], in [9], the optimal control of linear time-varying delay systems is given without using the solution of PDEs.

In [7], the linear quadratic suboptimal control is addressed for linear time-invariant systems with point-wise and distributed delays by using the solution of ordinary differential equation (ODE). Effort is also invested in inverse optimal control [8] for time-delay systems wherein one finds an optimal policy and then tries to identify a cost function that the policy minimizes. Note that the effort from [7, 8, 10] requires that the state vector and system matrices are known.

The output feedback control of time-delay systems is addressed in [2, 11–13]. The output feedback control in [11] mitigates disturbances for input delayed systems with uncertain dynamics. In contrast, a chain of predictors is employed in [12] to design the output feedback control of linear time-invariant systems with input, state and output delays. Moreover, the objective of output feedback control methods in [2, 14] is to stabilize the linear time-delayed systems under known system dynamics.

To overcome the need for system dynamics, either a learning mechanism [15] or an adaptive identifier based control [16] has to be employed. Adaptive control of continuous and discrete time-delay systems is addressed in [17–27]. In [19–21], adaptive controllers are proposed for time-delay systems with input delay using input-output representation. A predictor-based model reference adaptive control is utilized in [25] for a nonlinear time-delay system with known multiple state and input delays. However, the control policies [19–21, 25] are not optimal and require state vector.

In [17], the optimal adaptive output regulation of a linear discrete-time system with input delay is addressed using reinforcement learning (RL). In contrast, an adaptive identifier is studied in [18] for linear continuous time-delay systems under the assumption that the system is asymptotically stable. To our knowledge, the optimal adaptive control (OAC) of uncertain linear continuous-time systems with state and input delays using output feedback is not attempted in the literature.

Therefore, this paper aims to regulate a linear continuous-time system with known state and input delays and with known and uncertain dynamics. First, by using a cost function over infinite time horizon a value function is defined. By using the stationarity condition on the Hamiltonian or Bellman type equation (BTE) when the system dynamics are known, the optimal regulator using output feedback is determined by deriving an output delay algebraic Riccati equation (DARE).

Subsequently, for OAC, an adaptive identifier, which is introduced to estimate the system dynamics, is tuned by using output, input and their delayed values. Next, the adaptive identifier and certainty equivalence optimal control approach using the estimated DARE are considered together to design the optimal adaptive regulator. The boundedness of the output, state and parameter estimation errors are shown using Lyapunov theory provided the initial state of the observer, parameter and output estimation errors are in a compact set. In summary, the proposed approach does not require both the solution of PDE and full knowledge of the system dynamics while it is a memoryless solution.

Notations. Throughout the paper, $\|.\|$ and $\|.\|_F$ denote the Euclidean vector norm and Frobenius matrix norm, respectively, I_n is the $n \times n$ identity matrix, and 0_n is the $n \times n$ zero matrix. The controllability of the pair (A, B) is denoted as $ctrb(A, B)$. The rank of the matrix A is defined as $rank(A)$. The minimum and maximum singular values and eigen values of the matrix A are represented by $\delta_{min}(A)$, $\delta_{max}(A)$, $\lambda_{min}(A)$ and $\lambda_{max}(A)$, respectively. The L_2 norm for the function of time $f(t)$ is defined as $\|f\|_2 \triangleq \left(\int_0^\infty |f(\tau)|^2 d\tau\right)^{1/2}$ and we consider

$f \in L_2$ when $\|f\|_2$ exists or it is finite. The L_∞ norm is defined as $\|f\|_\infty \triangleq \sup_{t \geqslant 0} |f(t)|$, and we denote $f \in L_\infty$ when $\|f\|_\infty$ exists. The $diag\{A, B\}$ denotes a diagonal matrix of matrices A and B. The pseudo inverse of matrix A is defined as $A^\dagger = (A^T A)^{-1} A^T$.

2. PROBLEM STATEMENT AND OPTIMAL CONTROL

In this section, the problem formulation is introduced and the optimal control approach is presented.

2.1. PROBLEM STATEMENT

Consider a linear continuous-time system with state and input delays represented by

$$\begin{cases} \dot{x}(t) = A_0 x(t) + A_1 x(t - d_x) + B_0 u(t) + B_1 u(t - d_u) \\ y(t) = C x(t) \end{cases}, \tag{1}$$

with

$$\begin{cases} x(\theta) = \varphi_x(\theta), & \theta \in [-d_x, 0] \\ u(\theta) = \varphi_u(\theta), & \theta \in [-d_u, 0] \end{cases}, \tag{2}$$

where $x(t) \in \mathbb{R}^n$, $u(t) \in \mathbb{R}^m$ and $y(t) \in \mathbb{R}^p$, denote the state, control input and system output, with known constant state and input delays as d_x and d_u, respectively. The matrices $A_0 \in \mathbb{R}^{n \times n}$ and $A_1 \in \mathbb{R}^{n \times n}$, $B_0 \in \mathbb{R}^{n \times m}, B_1 \in \mathbb{R}^{n \times m}$, and $C \in \mathbb{R}^{p \times n}$ denote the drift dynamics, input matrices, and known output matrix, respectively. The history functions φ_x and φ_u are known and considered continuously differentiable.

Assumption 1. The linear continuous-time system (1) with state and input delays is considered controllable and observable. Matrix C satisfies $\lambda_{min}(C^T C) > 0$.

Remark 1. Controllability and observability properties ensure the existence of a solution to the DARE. Since $C^T C$ is symmetric, it will be positive definite [28], and pseudo inverse $C^\dagger = (C^T C)^{-1} C^T$ exists.

Next, consider the linear continuous-time system with state and input delays defined in (1). Design a control policy $u(t) = Ky(t)$, where $K \in \mathbb{R}^{m\times p}$ is a user defined gain matrix to be defined later, such that the closed-loop system is asymptotically stable, i.e. $\lim_{t\to\infty} x(t) \to 0$, while minimizing the quadratic cost function, $V(x)$, over infinite horizon given by

$$V(x(t)) = \int_t^{\infty} \left(y(\tau)^T \tilde{Q} y(\tau) + u(\tau)^T R u(\tau)\right) d\tau, \tag{3}$$

where $\tilde{Q} \in \mathbb{R}^{n\times n}$ and $R \in \mathbb{R}^{m\times m}$ represent user-defined positive definite matrices. Differentiating the value function (3), yields the following Hamiltonian function [29]

$$H(y, V, u) \equiv \frac{\partial V}{\partial t} + y^T \tilde{Q} y + u^T R u = 0, \tag{4}$$

where $\partial V/\partial t$ denotes the time derivative of the value function with respect to time. In [6, 7], the Hamiltonian equation (4) is referred to as the BTE. For the case of an uncertain linear continuous-time system with input and state delays, the objective is to design an optimal control policy in order to ensure that the closed-loop system is bounded while minimizing the value functional in (3).

2.2. OPTIMAL CONTROL WITH KNOWN DYNAMICS

An optimal output feedback approach is proposed in this section for the linear continuous-time system (1) with state and input delays. The BTE is defined using the value function. The optimal control policy is determined by using the stationarity condition on BTE provided the solution to a given DARE can be found. Finally, the asymptotic stability of the closed-loop system is shown. The following lemmas are required in order to proceed.

Lemma 1. Consider the linear continuous-time system (1) with state and input delays. Then, the derivative of the delayed input $u(t - d_u)$ along with the non-delay input $u(t)$, becomes $B_1 \partial u(t - d_u)/\partial u(t) = e^{A_0 d_u} B_1$.

Lemma 2. Let M_1 and N_1 be positive definite matrices, i.e. $v^T M_1 v > 0$ and $v^T N_1 v > 0$ for any vector $v \neq 0$. Then, $M_1 + N_1$ is a positive definite matrix [28].

To proceed, let the value function $V(y(t))$ be defined as

$$V(y) = y^T P_y y + \alpha_1 \int_{t-d_x}^{t} y^T(s)y(s)ds + \alpha_2 \int_{t-d_u}^{t} y^T(s)y(s)ds, \tag{5}$$

where $P_y \in \mathbb{R}^{n\times n}$ is a positive definite matrix and $\alpha_1 > 0$ and $\alpha_2 > 0$ are positive scalars. Now, using (5), the BTE (4) can be written as

$$\begin{aligned} BTE &= (A_0 x + A_1 x(t-d_x) + B_0 u(t) + B_1 u(t-d_u))^T C^T P_y y \\ &+ y^T P_y C(A_0 x + A_1 x(t-d_x) + B_0 u(t) + B_1 u(t-d_u)) \\ &+ \alpha_1 y^T y - \alpha_1 y^T (t-d_x) y(t-d_x) \\ &+ \alpha_2 y^T y - \alpha_2 y^T (t-d_u) y(t-d_u) + y^T \tilde{Q} y + u^T R u \end{aligned} \quad . \tag{6}$$

By applying the stationarity condition $\frac{\partial(BTE)}{\partial u} = \frac{2\partial u^T B_0^T C^T P_y y(t)}{\partial u} + \frac{2\partial u(t-d_u)^T B_1^T C^T P_y y(t)}{\partial u} + 2Ru = 0$ and using Lemma 1, one has $2B_0^T C^T P_y y(t) + 2(B_1 e^{A_0 d_u})^T C^T P_y y(t) + 2Ru = 0$ which results in

$$u^*(t) = -R^{-1} B^T C^T P_y y(t), \tag{7}$$

where $K = R^{-1} B^T C^T P_y$,$B = B_0 + e^{A_0 d_u} B_1$. The matrix $P_y \in \mathbb{R}^{n\times n}$ is obtained from the kernel matrix $P = C^T P_y C$ with $P \in \mathbb{R}^{n\times n}$ being the solution of the following output DARE that is derived in Theorem 1

$$A_0^T P + P A_0 - P B R^{-1} B^T P + \bar{Q} = 0, \tag{8}$$

with $R > 0$, $\bar{Q} = (\alpha_1 I + \alpha_2 I)C^T C + Q$, $Q \in \mathbb{R}^{n\times n}$ a user defined positive definite matrix. Since $\lambda_{min}(C^T C) > 0$, then, $C^T C$ becomes a positive definite (PD) matrix as for any vector v, one has $v^T C^T C v = (Cv)^T (Cv) > 0$. Moreover, $\alpha_1 > 0$, $\alpha_2 > 0$ and Q is also a PD matrix and, therefore, based on Lemma 2, $\bar{Q}$ becomes a PD matrix.

Remark 2. (Existence and uniqueness of the solution to the DARE). To guarantee the existence and uniqueness positive definite solution to (8), i.e. $P_y > 0$, the pairs (A_0, B) and $(A_0, \sqrt{\bar{Q}})$ must be controllable and observable, respectively. Based on Assumption 1, the pair (A_0, B) is controllable. On the other hand, observability of the system, i.e. $(A_0, \sqrt{C})$, results in the observability of $(A_0, \sqrt{\bar{Q}})$. Therefore, there exists a unique solution to the DARE (8) if Assumption 1 is fulfilled.

Remark 3. Note that it has been shown that the solution to (8) is a symmetric PD matrix, i.e. $P > 0$. Now, it is shown that matrix P_y is PD. To this end, from the definition of P_y one has $P_y = C^\dagger P (C^\dagger)^T$. Using the definition of positive semi-definiteness, let z_x be a vector, then, $z_x^T C^\dagger P (C^\dagger)^T z_x \geqslant 0$ can be written as $\bar{z}^T P \bar{z}$ where $\bar{z} = z_x C^\dagger$. Since $P_y > 0$, then, $\bar{z}^T P \bar{z} \geqslant 0$. Now, assume $\bar{z}^T P \bar{z} = 0$ and since $P > 0$, it implies that $\bar{z} = 0$, and consequently, $C^\dagger z_x = 0$. On the other hand, $C^\dagger$ is a full rank matrix which implies $ker(C^\dagger) = 0$. Finally, $C^\dagger z_x = 0$ concludes $z_x = 0$. This means $P_y > 0$ is PD.

Theorem 1. Consider the linear continuous-time system (1) with input and state delays. Let Assumption 1 be satisfied and the positive definite matrix P_y be the solution to the output DARE (8). Then, the control input (7) obtained using BTE is optimal and it makes the closed-loop system asymptotically stable if the design parameters are selected such that $\alpha_2 > \frac{\delta_{max}(PBR^{-1}B_1^T P)}{\lambda_{min}(C^T C)}$, $\alpha_1 > \frac{\delta_{max}(A_1^T P)}{\lambda_{min}(C^T C)}$, and $\lambda_{min}(Q) > \delta_{max}\left(PBR^{-1}(B - 2B_0)^T P\right) + \delta_{max}(A_1^T P) + \delta_{max}(PBR^{-1}B_1^T P)$.

Proof. Let the Lyapunov candidate function be defined as (5). Taking the derivative of (5) along the system dynamics (1) and considering $P = C^T P_y C$, one has $\dot{V} = x^T \left(A_0^T P + P A_0 + (\alpha_1 + \alpha_2) C^T C\right) x + 2x^T(t - d_x) A_1^T P x - \alpha_1 x^T (t - d_x) C^T C x(t - d_x) - \alpha_2 x^T (t - d_u) C^T C x(t - d_u) + 2u^T B_0^T P x + 2u^T (t - d_u) B_1^T P x$. Substituting the optimal control

(7) and using the output DARE (8), $\dot{V}$ becomes $\dot{V} = -x^T Q x + x^T PBR^{-1}(B - 2B_0)^T Px + 2x^T(t - d_x)A_1^T Px - \alpha_1 x^T(t - d_x)C^T Cx(t - d_x) + -2x^T(t - d_u)PBR^{-1}B_1^T Px - \alpha_2 x^T(t - d_u)C^T Cx(t - d_u)$. Next, utilizing the Young's inequality, and considering the condition $\lambda_{\min}(C^T C) > 0$, one has

$$
\begin{aligned}
\dot{V} \leqslant & - \left(\lambda_{\min}(Q) - \delta_{max}\left(A_1^T P\right) - \delta_{max}(PBR^{-1}B_1^T P)\right. \\
& \left. - \delta_{max}(PBR^{-1}(B - 2B_0)^T P)\right)\|x\|^2 \\
& - (\alpha_1 \lambda_{\min}(C^T C) - \delta_{max}(A_1^T P))\|x(t - d_x)\|^2 \\
& - (\alpha_2 \lambda_{\min}(C^T C) - \delta_{max}(PBR^{-1}B_1^T P))\,\|x(t - d_u)\|^2
\end{aligned}
\tag{9}
$$

Therefore, one can conclude that if conditions in the statement of the theorem are met, $\dot{V} < 0$. This states that the optimal control input (7), guarantees the asymptotic stability of the closed-loop system (1). ■

Remark 4. The design variable selection in the above theorem requires the solution to the DARE (8), which in turn needs system matrices. By slightly altering the proof, one can derive a different form of DARE, whereas the one in (8) becomes a traditional ARE when the delays are set to zero. For instance, when $A_1 = B_1 = 0$ and the the second and third terms in the value function given in (5) are zero due to $\alpha_1 = \alpha_2 = 0$, the output DARE (8) leads to the traditional algebraic Riccati equation with $\tilde{Q} = Q = C^T C$ and $B = B_0$ when, the state vector is measurable.

Remark 5. Note that the obtained DARE (8) depends upon the coefficient matrix of the input delay and design variable α_1, which is selected using the coefficient matrix of the state delay. Also, the control policy is memoryless and it does not need the solution to a PDE.

3. OPTIMAL ADAPTIVE OUTPUT FEEDBACK CONTROL

In this section, OAC using output feedback is proposed for linear time-delay systems (1) with input and state delays in the presence of uncertain dynamics. Both traditional adaptive control and integral reinforcement learning (IRL) methods can be used. For adaptive control, an adaptive identifier is employed to estimate the dynamics, and by applying the principle of certainty equivalence, a controller is designed using both the estimated parameters and the solution to the estimated ARE. In contrast, in the IRL, the solution to the ARE is found through adaptive estimation of value function (5). In our preliminary version [30], the value function is estimated and the Kernel matrix P_y^*, is generated for time-delay systems with state delay.

By contrast, in this paper, a certainty equivalence based optimal adaptive framework will be introduced. First, an estimated control input and DARE are presented using estimated dynamics. Then, an adaptive observer is designed to estimate the system dynamics with the parameters tuned by measured output. Closed-loop stability is demonstrated for ideal and general cases. It is shown that for the ideal case, the output and the state, observer estimation error and the parameter estimation errors converge to zero asymptotically provided a Persistency of Excitation (PE) condition on the estimated control input is satisfied.

Then, in the general case, by adding an extra term to the parameter update law, the robust optimal adaptive control performance, in terms of boundedness of the overall closed-loop system, is presented using Lyapunov theory. The overall structure of the proposed optimal adaptive control is shown in Figure 1. The following assumption is needed in order to proceed.

Assumption 2. The system matrices of the linear continuous-time system (1) with state and input delays are considered uncertain, whereas their upper bounds are known, i.e. $\|A_0\| \leqslant A_{0M}$, $\|A_1\| \leqslant A_{1M}$, $\|B_0\| \leqslant B_{0M}$, $\|B_1\| \leqslant B_{1M}$. Alternatively, a nominal model [16] is available.

Remark 6. The bounds on the uncertain system matrices given in the Assumption 2 can be employed to determine the solution to DARE for design variable selection. Alternatively, it is typical to have an apriori information such as a nominal model [16, 31] of the linear system for this purpose. However, this information is not needed to generate the optimal control policy.

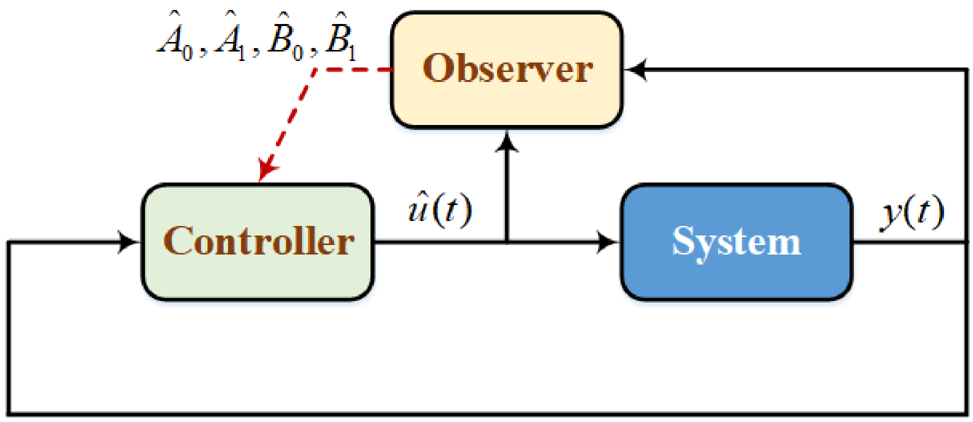


Figure 1. The proposed optimal adaptive control approach.

From Figure 1, it can be observed that the control input applied to the system has only access to the estimated parameters provided by the adaptive observer due to the uncertainty in the system dynamics. Therefore, using the principle of certainty equivalence [16], the estimated optimal control input can be expressed by

$$\hat{u}(t) = -R^{-1}\hat{B}^T C^T \hat{P}_y y(t), \tag{10}$$

where $\hat{B} = \hat{B}_0 + e^{\hat{A}_0 d_u}\hat{B}_1$ with the matrix $\hat{P}_y$ being the solution of $\hat{P} = C^T\hat{P}_y C$, where the kernel matrix $\hat{P}$ is the solution to the following estimated DARE given by

$$\hat{A}_0^T\hat{P} + \hat{P}\hat{A}_0 - \hat{P}\hat{B}R^{-1}\hat{B}^T\hat{P} + \bar{Q} = 0, \tag{11}$$

with $\bar{Q} = (\alpha_1 I + \alpha_2 I)C^T C + Q$, $Q \in \mathbb{R}^{n\times n}$ a user defined positive definite matrix.

The estimated DARE needs to be solved at each time instant by using the online estimated pair $(\hat{A}_0, \hat{B})$ of the linear time-delayed system similar to the case of non-delay system [16]. For the solution, $\hat{P}$, to exist, the pair $(\hat{A}_0, \hat{B})$ has to be stabilizable [16] and observable at each time. Observability of the system, i.e. $(\hat{A}_0, \sqrt{C})$, results in the observability of $(\hat{A}_0, \sqrt{\bar{Q}})$. This ensures that there exists a unique solution, $\hat{P}$, to the output DARE (11).

3.1. IDEAL CASE

In this subsection, after presenting a novel adaptive observer, the state estimation error is defined by using the estimated and actual state vector. The dynamics of the observer estimation error are derived and an update law for the unknown parameters is introduced. Finally, under mild conditions, the Lyapunov stability analysis is employed to show that the system output and observer state errors converge to zero asymptotically. Next under a PE condition [32], it is shown that the parameter estimation error converges to zero. Then, the estimated control input attains the optimal control policy.

By using the system dynamics (1), the adaptive observer is constructed as

$$\begin{cases} \dot{\hat{x}}(t) = A_m C^{\dagger}(\hat{y}(t) - y(t)) + \hat{B}_0 u(t) \\ +A_m C^{\dagger}(\hat{y}(t-d_x) - y(t-d_x))) + \hat{B}_1 u(t-d_u) \\ +\hat{A}_0 C^{\dagger} y(t) + \hat{A}_1 C^{\dagger} y(t-d_x) \\ \hat{y}(t) = C\hat{x}(t) \end{cases}, \tag{12}$$

where $\hat{x} \in \mathbb{R}^{n\times 1}$ and $\hat{y} \in \mathbb{R}^{p\times 1}$ denote the observer state and estimated output, respectively. Moreover, $A_m \in \mathbb{R}^{n\times n}$ is a design gain matrix. The matrices $\hat{A}_0 \in \mathbb{R}^{n\times n}$, $\hat{A}_1 \in \mathbb{R}^{n\times n}$, $\hat{B}_0 \in \mathbb{R}^{m\times n}$ and $\hat{B}_1 \in \mathbb{R}^{m\times n}$ represent an estimate of the actual system dynamics A_0, A_1, B_0

and B_1, respectively. Denote the observer state estimation error as

$$e(t) = x(t) - \hat{x}(t). \tag{13}$$

Then, the dynamic of the observer state estimation error becomes

$$\begin{aligned} \dot{e}(t) &= A_m e(t) + A_m e(t - d_x) - \tilde{A}_0 C^{\dagger} y(t) \\ &\quad - \tilde{A}_1 C^{\dagger} y(t - d_x) - \tilde{B}_0 u(t) - \tilde{B}_1 u(t - d_u) \end{aligned}, \tag{14}$$

where parameter estimation error matrices are given by $\tilde{A}_0 = \hat{A}_0 - A_0$, $\tilde{A}_1 = \hat{A}_1 - A_1$, $\tilde{B}_0 = \hat{B}_0 - B_0$ and $\tilde{B}_1 = \hat{B}_1 - B_1$. It is important to note that the observer state estimate error is driven by parameter estimation error. The gain matrix A_m is selected such that A_m is Hurwitz and, therefore, there exists a positive definite matrix $P_L \in \mathbb{R}^{n\times n}$ as a solution to the following modified Lyapunov equation

$$(A_m)^T P_L + P_L(A_m) + \breve{Q} = 0\,, \tag{15}$$

where $\breve{Q} = Q_L + \alpha_L I_n$ is a design positive definite matrix, with $Q_L \in \mathbb{R}^{n\times n}$ as a positive definite design matrix and $\alpha_L > 0$ is a scalar. Now, define the parameter tuning law as

$$\begin{cases} \dot{\hat{A}}_0 = \gamma_0 C^{\dagger} \tilde{y}(t) \left(C^{\dagger} y(t)\right)^T \\ \dot{\hat{A}}_1 = \gamma_1 C^{\dagger} \tilde{y}(t) \left(C^{\dagger} y(t - d_x)\right)^T \\ \dot{\hat{B}}_0 = \beta_0 C^{\dagger} \tilde{y}(t) \hat{u}^T(t) \\ \dot{\hat{B}}_1 = \beta_1 C^{\dagger} \tilde{y}(t) \hat{u}^T(t - d_u) \end{cases}, \tag{16}$$

where $\tilde{y} = y - \hat{y}$ is the output estimation error and γ_0, γ_1, β_0 and β_1 represent positive adaptation gains. Define the matrix of all uncertain system parameters as $\hat{M} = diag\{\hat{A}_0, \hat{A}_1, \hat{B}_0, \hat{B}_1\}$. Then, the overall parameter estimation error can be defined as $\tilde{M} = \hat{M} - M = diag\{\tilde{A}_0, \tilde{A}_1, \tilde{B}_0, \tilde{B}_1\}$, where $M = diag\{A_0, A_1, B_0, B_1\}$. Using the bounds on the system dynamics and input matrices, one has $\|M\| \leqslant M_{max}$.

The parameter update law (16) uses the output, delayed output, input and delayed input to estimate the parameter matrices and, therefore, it is not a standard from of traditional adaptive control literature. Next, in the following theorem the performance of the optimal adaptive regulator under ideal conditions is demonstrated.

Theorem 2. (Ideal Case): Consider the linear system (1) with state and input delays and the observer dynamics (12). Let Assumptions 1 and 2 hold, and stabilizability of the pair $(\hat{A}_0, \hat{B})$ be satisfied with the overall closed-loop system equilibrium state being $(x = 0,\ e = 0,\ \tilde{M} = 0)$. Assume that the system is initialized in $(x(0), e(0), \hat{M}(0)) \in \{\Omega_x, \Omega_e, \Omega_M\}$ where Ω_x, Ω_e and Ω_M, are compact sets corresponding to system state, observer estimation error and parameters estimate, respectively. Let the estimated control input (10) obtained from the solution of estimated DARE (11) be persistently exciting and the parameter update law (16) be employed with P_L as the solution to the Lyapunov equation (15). Select the design parameters, θ_1, θ_2, θ_3, μ_1, and μ_2, as positive scalars that are defined in the proof. Then,

1. the observer state and output estimation errors along with state and output of the system converge to zero asymptotically;

2. under the PE condition (see Theorem 5.2.3 of [16]) on the estimated control input, the parameter estimation errors converge to zero and the estimated control input attains the optimal value defined by (7).

Proof. Define the Lyapunov candidate function as $V = V_1 + V_2$ where V_1 is considered to be the value function defined in (5) and

$$
\begin{aligned}
V_2 = e^T P_L e + \alpha_L \int_{t-d_x}^{t} e^T(s)e(s)ds \\
+ trace\left\{\frac{\tilde{A}_0^T P_L \tilde{A}_0}{\gamma_0}\right\} + trace\left\{\frac{\tilde{A}_1^T P_L \tilde{A}_1}{\gamma_1}\right\} \\
+ trace\left\{\frac{\tilde{B}_0^T P_L \tilde{B}_0}{\beta_0}\right\} + trace\left\{\frac{\tilde{B}_1^T P_L \tilde{B}_1}{\beta_1}\right\}
\end{aligned}
\tag{17}
$$

The first derivative of the overall Lyapunov function candidate becomes $\dot{V} = \dot{V}_1 + \dot{V}_2$. Then, the first derivative of V_1 using the estimated control input (10) and $P = C^T P_y C$ can be written as

$$
\begin{aligned}
\dot{V}_1 = (A_0 x + A_1 x(t-d_x) + B_0 \hat{u}(t) + B_1 \hat{u}(t-d_u))^T Px \\
+ x^T P(A_0 x + A_1 x(t-d_x) + B_0 \hat{u}(t) \\
+ B_1 \hat{u}(t-d_u)) + \alpha_1 x^T(t) C^T C x(t) \\
- \alpha_1 x^T(t-d_x) C^T C x(t-d_x) \\
+ \alpha_2 x^T C^T C x - \alpha_2 x^T(t-d_u) C^T C x(t-d_u)
\end{aligned}
\tag{18}
$$

Let the control input error be defined as $\tilde{u}(t) = \hat{u}(t) - u^*(t)$ where $u^*(t)$ is the actual optimal control input. Define $\tilde{\Gamma} = \hat{B}\hat{P} - B^T P$, then, $\|x^T\tilde{\Gamma}\|^2 \leqslant c_1\|x(t)\|^2 + c_2\|e(t)\|^2$ and $\|x(t-d_u)^T\tilde{\Gamma}\|^2 \leqslant c_3\|x(t-d_u)\|^2 + c_4\|e(t)\|^2$ where c_1, c_2, c_3, and c_4 are positive scalars. Then, one has $\hat{u} = -R^{-1}\hat{B}^T C^T \hat{P}_y y = -R^{-1}\hat{B}^T \hat{P} x = -R^{-1} B^T P x - R^{-1}\tilde{\Gamma}^T x$ where $\tilde{\Gamma} = \hat{B}\hat{P} - B^T P$. Using the estimated control input in terms of the actual control input and the control input error and after some simplifications and using $B_0 = B - e^{A_0 d_u} B_1$, using the DARE (8) and Young's inequality for the cross products, (18) becomes

$$
\begin{aligned}
\dot{V}_1 \leqslant & -\lambda_{\min}(Q)\|x\|^2 + \delta_{\max}(A_1^T P)\|x\|^2 + \delta_{max}(PBR^{-1}(B-2B_0)^T P)\|x\|^2 \\
& + \delta_{\max}(PBR^{-1}B_1^T P)\|x\|^2 + \delta_{\max}(A_1^T P)\|x(t-d_x)\|^2 \\
& - \alpha_1 \lambda_{min}(C^T C)\|x(t-d_x)\|^2 + \delta_{\max}(PBR^{-1}B_1^T P)\|x(t-d_u)\|^2 \\
& - \alpha_2 \lambda_{min}(C^T C)\|x(t-d_u)\|^2 + \delta_{\max}(R^{-1}B_1^T P)\|x\|^2 \\
& + c_3 \delta_{\max}(R^{-1}B_1^T P)\|x(t-d_u)\|^2 + (c_1+1)\delta_{\max}(R^{-1}B_0^T P)\|x\|^2 \\
& + (c_2 \delta_{\max}(R^{-1}B_0^T P) + c_4 \delta_{\max}(R^{-1}B_1^T P))\|e\|^2
\end{aligned}
\tag{19}
$$

Now, taking the first derivative of (17), using (14) and after a few manipulations, one has $\dot{V}_2 = e^T(t)\left((A_m)^T P_L + P_L(A_m) + \alpha_L I_n\right)e(t) + 2e^T(t-d_x)(A_m)^T P_L e(t) - \alpha_L e^T(t-d_x)e(t-d_x) - 2e^T(t)P_L\tilde{A}_0 C^\dagger y(t) - 2e^T(t)P_L\tilde{A}_1 C^\dagger y(t-d_x) - 2e^T(t)P_L\tilde{B}_0\hat{u}(t) - 2e^T(t)P_L\tilde{B}_1\hat{u}(t-d_u) + 2trace\{\tilde{A}_0^T P_L \dot{\tilde{A}}_0/\gamma_0\} + 2trace\{\tilde{A}_1^T P_L \dot{\tilde{A}}_1/\gamma_1\} + 2trace\{\tilde{B}_0^T P_L \dot{\tilde{B}}_0/\beta_0\} + 2trace\{\tilde{B}_1^T P_L \dot{\tilde{B}}_1/\beta_1\}$ which, by using properties of the trace function as $x^T y = trace(xy^T)$ and $trace(ABC) = trace(CAB) = trace(BCA)$, yields

$$
\begin{aligned}
\dot{V}_2 = & \, e^T(t)\left((A_m)^T P + P_L(A_m) + \alpha_L I_n\right)e(t) \\
& + 2e^T(t)P_L(A_m)e(t-d_x) - \alpha_L e^T(t-d_x)e(t-d_x) \\
& + 2trace\left\{\frac{\tilde{A}_0^T P_L \dot{\tilde{A}}_0}{\gamma_0} - \tilde{A}_0^T P_L (C^\dagger \tilde{y}(t))(C^\dagger y(t))^T\right\} \\
& + 2trace\left\{\frac{\tilde{A}_1^T P_L \dot{\tilde{A}}_1}{\gamma_1} - \tilde{A}_1^T P_L (C^\dagger \tilde{y}(t))(C^\dagger y(t-d_x))^T\right\} \\
& + 2trace\left\{\frac{\tilde{B}_0^T P_L \dot{\tilde{B}}_0}{\beta_0} - \tilde{B}_0^T P_L (C^\dagger \tilde{y}(t))\hat{u}^T(t)\right\} \\
& + 2trace\left\{\frac{\tilde{B}_1^T P_L \dot{\tilde{B}}_1}{\beta_1} - \tilde{B}_1^T P_L (C^\dagger \tilde{y}(t))\hat{u}^T(t-d_u)\right\}
\end{aligned}
$$

which, by substituting the update law (16), gives

$$\begin{aligned}\dot{V}_2 &= e^T(t)\left((A_m)^T P_L + P_L(A_m) + \alpha_L I_n\right)\\ &e(t) + 2e^T(t)P_L(A_m)e(t-d_x) - \alpha_L e^T(t-d_x)e(t-d_x)\end{aligned} \tag{20}$$

Using the modified Lyapunov equation (15) and Young's inequality, (20) becomes

$$\begin{aligned}\dot{V}_2 \leqslant &-(\lambda_{\min}(Q_L) - \delta_{\max}((A_m)^T P_L))\|e(t)\|^2\\ &-(\alpha_L - \delta_{\max}((A_m)^T P_L))\|e(t-d_x)\|^2\end{aligned} \tag{21}$$

Collecting the first derivatives of the Lyapunov function candidates V_1 and V_2, and after some simplifications by assuming $\|x\| = \|C^\dagger y\|$, the derivative of the overall Lyapunov candidate function, i.e. $\dot{V}$, becomes

$$\begin{aligned}\dot{V} \leqslant -\Big(&\theta_1\|C^\dagger y\|^2 + \theta_2\|C^\dagger y(t-d_x)\|^2 + \mu_1\|e(t)\|^2\\ &+\theta_3\|C^\dagger y(t-d_u)\|^2 + \mu_2\|e(t-d_x)\|^2\Big)\end{aligned} \tag{22}$$

In (22), $\theta_1 = \lambda_{\min}(Q) - \delta_{\max}(A_1^T P) - \delta_{max}(PBR^{-1}(B-2B_0)^T P) - \delta_{\max}(PBR^{-1}B_1^T P) - \delta_{\max}(R^{-1}B_1^T P) - (c_1+1)\delta_{\max}(R^{-1}B_0^T P)$, $\theta_2 = \alpha_1\lambda_{min}(C^T C) - \delta_{\max}(A_1^T P)$, $\theta_3 = \alpha_2\lambda_{min}(C^T C) - \delta_{\max}(PBR^{-1}B_1^T P) - c_3\delta_{\max}(R^{-1}B_1^T P)$, $\mu_1 = \lambda_{\min}(Q_L) - \delta_{\max}((A_m)^T P_L) - c_2\delta_{\max}(R^{-1}B_0^T P) - c_4\delta_{\max}(R^{-1}B_1^T P)$, and $\mu_2 = \alpha_L - \delta_{\max}((A_m)^T P_L)$.

i) This concludes that if the conditions in the statement of Theorem 2 are satisfied, $\dot{V} \leqslant 0$ and results in $x, x(t-d_x), x(t-d_u), e, e(t-d_x) \in L_2$ since, C is a constant matrix and $y, y(t-d_x), y(t-d_u) \in L_2$. The boundedness of x, y and e imply that the estimated parameters (16) and the estimated control input $\hat{u}$ become bounded. Then, considering the observer estimation error dynamics (14), one has $\dot{e} \in L_2$ and according to Lemma 3.2.5 in [16] one has $e \to 0$ when $t \to \infty$. The same argument can be used for the system state x and, consequently, the output of the system.

ii) Next, as shown (see Theorem 5.2.3 of [16], [32]) when the PE condition is satisfied, the parameter estimation error converges to zero asymptotically. According to the definition of $\tilde{u}(t) = \hat{u}(t) - u^*(t) = -R^{-1}\tilde{\Gamma}Cx$ with $\tilde{\Gamma} = \hat{B}\hat{P} - B^T P$, when the estimated parameters converge to the actual parameters, the estimated DARE will become actual DARE since the coefficients of the quadratic equations become equal. This implies that the estimated $\hat{P}$ converges to P and, therefore, $\tilde{\Gamma} \to 0$. This further implies that the control input error converges to zero, eventually, which results in the convergence of the estimated control input to the optimal one. ■

Remark 7. The design parameters provided in the statement of Theorem 2 can be determined using Assumption 2. The parameter update law (16) depends upon the measured system output and control input. Asides from this update law (16), other approaches, such as a projection approach can also be employed.

Remark 8. The PE condition helps to ensure the convergence of the parameter estimation errors [16] and actual control input to its optimal value. A small random noise [33] can be utilized for PE. Alternatively, in [34], the PE condition can be relaxed by modifying the parameter tuning law as given in the composite model reference adaptive control.

3.2. GENERAL CASE

In this subsection, by using the adaptive observer from the previous subsection, modified update laws for the unknown parameters to mitigate parameter drift are introduced. The modified update laws help in the boundedness of the closed-loop system even in the presence of disturbances and ummodelled dynamics, if any. Finally, under mild conditions, Lyapunov stability analysis is employed to show the boundedness of the closed-loop system.

Consider modified parameter update law given by

$$\begin{cases} \dot{\hat{A}}_0 = \gamma_0 C^\dagger \tilde{y}(t) \left(C^\dagger y(t)\right)^T - \gamma_0 \hat{A}_0 \\ \dot{\hat{A}}_1 = \gamma_1 C^\dagger \tilde{y} \left(C^\dagger y(t-d_x)\right)^T - \gamma_1 \hat{A}_1 \\ \dot{\hat{B}}_0 = \beta_0 C^\dagger \tilde{y} \hat{u}^T(t) - \beta_0 \hat{B}_0 \\ \dot{\hat{B}}_1 = \beta_1 C^\dagger \tilde{y} \hat{u}^T(t-d_u) - \beta_1 \hat{B}_1 \end{cases}, \tag{23}$$

where the adaptation gains γ_0, γ_1, β_0 and β_1 are positive design scalars. The above parameter update is considered to be robust and mitigates the drift. Given the boundedness of the parameter estimation errors, it can be concluded that the estimated control input is bounded close to its optimal value.

Theorem 3. (General Case): Consider the linear continuous time system (1) with input and state delays. Let Assumptions 1 and 2 and stabilizability of the pair $(\hat{A}_0, \hat{B})$ be satisfied with the overall closed-loop system equilibrium state being $(x = 0,\ e = 0,\ \tilde{M} = 0)$. Assume that the system is initialized in $(x(0), e(0), \hat{M}(0)) \in \{\Omega_x, \Omega_e, \Omega_M\}$, where Ω_x, Ω_e and Ω_M are compact sets corresponding to system state, observer estimation error and parameters estimate, respectively. Let the estimated control input (10) obtained from the solution of estimated DARE (11) be persistently exciting and parameters tuning law (23) be utilized. Consider the design parameters, θ_1, θ_2, θ_3, θ_4, θ_5,θ_6, μ_1, and μ_2 as positive scalars which are defined in the proof. Then,

1. the overall closed-loop system becomes bounded; and
2. the estimated control input will be bounded close to its optimal value as defined by (7).

Proof. Consider the overall Lyapunov function $V = V_1 + V_2$ with V_1 and V_2 defined in Theorem 2. Taking the first derivative of V_1 and performing the same procedure and simplification given in the proof of Theorem 2, and assuming $\|x^T \tilde{\Gamma}\|^2 \leqslant c_0 + c_1\|x\|^2 + c_2\|\tilde{M}\|^2$

and $\|x(t-d_u)^T\tilde{\Gamma}\|^2 \leqslant c_3 + c_4\|x(t-d_u)\|^2 + c_5\|\tilde{M}\|^2$ where c_0, c_1, c_2, c_3, c_4 and c_5 are positive scalars, one has

$$
\begin{aligned}
\dot{V}_1 \leqslant\ & -\lambda_{\min}(Q)\|x\|^2 + \delta_{\max}(A_1^T P)\|x\|^2 \\
& + \delta_{max}(PBR^{-1}(B-2B_0)^T P)\|x\|^2 + \delta_{\max}(PBR^{-1}B_1^T P)\|x\|^2 \\
& + \delta_{\max}(A_1^T P)\|x(t-d_x)\|^2 - \alpha_1\lambda_{min}(C^T C)\|x(t-d_x)\|^2 \\
& + \delta_{\max}(PBR^{-1}B_1^T P)\|x(t-d_u)\|^2 - \alpha_2\lambda_{min}(C^T C)\|x(t-d_u)\|^2 \\
& + c_4\delta_{\max}(R^{-1}B_1^T P)\|x(t-d_u)\|^2 \\
& + \delta_{\max}(R^{-1}B_0^T P)\|x\|^2 + \delta_{\max}(R^{-1}B_1^T P)\|x\|^2 \\
& + c_1\delta_{\max}(R^{-1}B_0^T P)\|x\|^2 + c_2\delta_{\max}(R^{-1}B_0^T P)\|\tilde{M}\|^2 \\
& + c_4\delta_{\max}(R^{-1}B_1^T P)\|x(t-d_u)\|^2 + c_5\delta_{\max}(R^{-1}B_1^T P)\|\tilde{M}\|^2 + c_0 + c_3
\end{aligned}
\tag{24}
$$

The first derivative of V_2 by carrying out the same steps provided in Theorem 2 with the parameter update law (23) and, after some simplifications, gives

$$
\begin{aligned}
\dot{V}_2 \leqslant -\Big[& \mu_1\|e(t)\|^2 + \mu_2\|e(t-d_x)\|^2 \\
& +2\delta_{\max}(P_L)\left(\|\tilde{M}\| - \frac{M_{max}}{2}\right)^2 - \mu_3\Big]
\end{aligned}
\tag{25}
$$

with $\mu_1 = \lambda_{\min}(Q_L) - \delta_{\max}((A_m)^T P_L)$, $\mu_2 = \alpha_L - \delta_{\max}((A_m)^T P_L)$ and $\mu_3 = \delta_{\max}(P_L)M_{max}^2/2$, where P_L and Q_L are defined in (15).

Now, collecting the first derivatives of the Lyapunov function candidates V_1 and V_2, and using $\|x\| = \|C^\dagger y\|$, the derivative of the overall Lyapunov candidate function $\dot{V}$ becomes

$$
\begin{aligned}
\dot{V} \leqslant -\Big(& \theta_1\|C^\dagger y\|^2 + \theta_2\|C^\dagger y(t-d_x)\|^2 \\
& + \theta_3\|C^\dagger y(t-d_u)\|^2 + \mu_1\|e(t)\|^2 + \mu_2\|e(t-d_x)\|^2 \\
& +\theta_4\big(\|\tilde{M}\| - \theta_5\big)^2 - \theta_6\Big)
\end{aligned}
\tag{26}
$$

where $\theta_1 = \lambda_{\min}(Q) - \delta_{\max}(A_1^T P) - \delta_{\max}(PBR^{-1}B_1^T P) - \delta_{\max}(R^{-1}B_1^T P) - (c_1+1)\delta_{\max}(R^{-1}B_0^T P) - \delta_{max}(PBR^{-1}(B-2B_0)^T P)$, $\theta_2 = \alpha_1\lambda_{min}(C^T C) - \delta_{\max}(A_1^T P)$, $\theta_3 = \alpha_2\lambda_{min}(C^T C) - \delta_{\max}(PBR^{-1}B_1^T P) - c_4\delta_{\max}(R^{-1}B_1^T P)$, $\theta_4 = 2\delta_{\max}(P_L) - c_5\delta_{\max}(R^{-1}B_1^T P) - c_2\delta_{\max}(R^{-1}B_0^T P)$, $\theta_5 = (\delta_{max}(P_L)M_{max})/\theta_4$ and $\theta_6 = (\delta_{max}^2(P_L)M_{max}^2)/\theta_4 + c_0 + c_3$.

i) This concludes that $\dot{V} < 0$ provided the design variables are selected as discussed in the statement of Theorem 3 and when any one of the bounds hold: $\|y\| > (1/\|C^\dagger\|)\sqrt{\theta_6/\theta_1}$ or $\|y(t-d_x)\| > (1/\|C^\dagger\|)\sqrt{\theta_6/\theta_2}$ or $\|y(t-d_u)\| > (1/\|C^\dagger\|)\sqrt{\theta_6/\theta_3}$ or $\|e(t)\| > \sqrt{\theta_6/\mu_1}$ or $\|e(t-d_x)\| > \sqrt{\theta_6/\mu_2}$ or $\|\tilde{M}\| > \theta_5 + \sqrt{\theta_6/\theta_4}$. This implies the boundedness of the overall closed-loop system.

ii) Next, according to the definition of $\tilde{u}(t) = \hat{u}(t) - u^*(t) = -R^{-1}\tilde{\Gamma}y$ with $\tilde{\Gamma} = \hat{B}\hat{P} - B^T P$. Since the estimated parameters are bounded to the actual parameters, the solution to the estimated DARE will be close and bounded to the actual DARE, because the DARE is a quadratic equation as their coefficients between estimated and actual equations are bounded. Therefore, $\tilde{\Gamma}$ is bounded. This makes the estimated control input bounded to the optimal control input, and the bound can be made small by selecting the design parameters appropriately. ■

Remark 9. The bounds for the parameter estimation error and control policy can be adjusted via design parameter selection to ensure that the estimated control input approaches the optimal value.

Remark 10. The optimal adaptive output feedback control of time-delay systems presented in this paper can be extended using state feedback by taking $C = I_n$. Then, the optimal feedback becomes $u^*(t) = Kx(t)$ with the value function defined with respect to the system state vector $x(t)$. Moreover, instead of an adaptive observer, an adaptive estimator can be utilized for the system dynamics.

In the next section, simulation results are provided to confirm the effectiveness of the proposed approach.

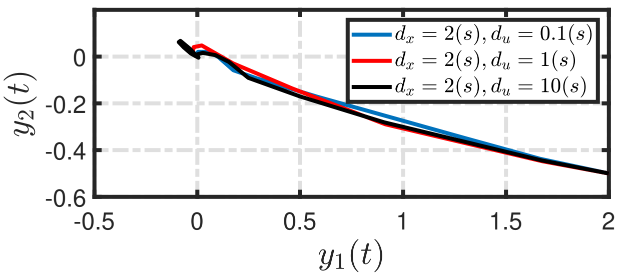


Figure 2. The output trajectories of the system (27) for different values of the state and input delays.

4. SIMULATION RESULTS

In this section, an uncertain linear continuous-time system with state and input delays is considered. Then, using the estimated optimal policy (10), the estimated DARE (11) and the parameter update law (23), the validity of near optimal adaptive approach is verified.

Example. Consider the linear continuous-time system with state and input delays given by

$$\begin{cases} \dot{x}(t) = A_0 x(t) + A_1 x(t-1) + B_0 u(t) + B_1 u(t-1) \\ y(t) = Cx(t) \end{cases}, \tag{27}$$

where

$$A_0 = \begin{bmatrix} 1 & 4 \\ 1 & -5 \end{bmatrix}, A_1 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

$$B_0 = \begin{bmatrix} 10 \\ 0 \end{bmatrix}, B_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, C = \begin{bmatrix} 1 & 0 \\ 0.8 & 1 \end{bmatrix}.$$

The history profiles for state and input are considered as $\varphi_x(\theta) = [1, 1]^T, \theta \in [-1, 0]$ and $\varphi_u(\theta) = 0, \theta \in [-1, 0]$. Next, in order to determine the design parameters as needed for the optimal adaptive regulator, the bounds in Assumption 2 (or nominal model of known dynamics) are considered. In addition, the following design parameters are selected: $Q =$

$10I_2$ and $R = 1$, $\alpha_1 = 1$ and $\alpha_2 = 3$ which results in $\tilde{Q} = 4(C^TC) + Q = \begin{bmatrix} 15.4 & 2.7 \\ 2.7 & 13.3 \end{bmatrix}$. Then, the solution to the DARE (8) becomes

$$P^* = \begin{bmatrix} 0.28 & 0.03 \\ 0.03 & 1.25 \end{bmatrix}$$

One can compute $\alpha_1 = 1 > (\delta_{max}(A_1P)/\lambda_{min}(C^TC)) = 0.62$ and $\alpha_2 = 3 > (\delta_{max}(PB_yR_{-1}B_1^TP)/\lambda_{min}(C^TC)) = 2.7$ and $\lambda_{min}(Q) = 10 > \delta_{max}\left(PBR^{-1}(B-2B_0)^TP\right) + \delta_{max}(A_1^TP) + \delta_{max}(PBR^{-1}B_1^TP) = 8.8$.

In order to show the stability of the overall system with delays, the output trajectory of the system for different values of state and input delays under known dynamics is provided in Figure 2. One can see that delay makes the convergence time longer, yet the system is stable. Now, for the optimal adaptive regulator design, the design matrix is defined as as $A_m = \begin{bmatrix} -2 & 0 \\ -1.6 & -2 \end{bmatrix}$ which results in the Hurwitz matrix A_m. The initial conditions for the estimation parameters are selected as $\hat{A}_0(0) = \begin{bmatrix} 0 & 0.5 \\ -0.5 & 0.5 \end{bmatrix}$, $\hat{A}_1(0) = \begin{bmatrix} 0 & 0.5 \\ 0.5 & 0.5 \end{bmatrix}$, $\hat{B}_0(0) = \begin{bmatrix} 0.1 \\ 0.1 \end{bmatrix}$ and $\hat{B}_1(0) = \begin{bmatrix} 0.1 \\ 0.1 \end{bmatrix}$. All other design variables are selected from known upper bounds as per Assumption 2.

The system output, output estimation error and the control input are depicted in Figure 3. A persistently exciting noise is added to the control input when $t < 20$ to ensure the parameter convergence before the system output converges. After the persistently exciting noise is released, the output converges and bounded near zero. The control input can be observed in Figure 3c, which is also bounded. One can see that, after the persistently exciting noise is removed, the control input tends to the value in which output of the system becomes bounded. The error in the solution of the estimated DARE and actual DARE is plotted in Figure3d, which indicates that it is bounded close to the optimal solution, thus, confirming the optimal nature of the adaptive regulator.

Next, the parameter estimation errors are shown in Figure 4. The performance is shown to be acceptable by using the update law (23), the estimated control input (10) and the estimated DARE (11). According to Theorem 4, the observer and the parameter estimation errors must remain bounded, which is confirmed in Figure 4. From the results, it is clear that the parameter and output estimation errors converge at the same time when the PE condition is removed. All results presented in Figure 3 confirm the theoretical claims in the paper.

5. CONCLUSION

In this paper, the near optimal adaptive regulation of uncertain linear continuous-time systems with state and input delays by using state and output feedback is addressed under mild assumption that the system matrices are uncertain, whereas, their bounds are known. First, the control policy for the linear system with state and input delays with known dynamics is shown to be optimal and renders the system asymptotically stable by selecting a value functional and Hamiltonian (or BTE), and a solution to the DARE is found. The PE condition ensured convergence of the parameter estimation error.

Next, for the case of uncertain dynamics, an adaptive identifier is utilized with estimated DARE and control policy to show that the overall closed-loop system is bounded in the general case. The estimated optimal control is shown to approach the actual optimal control input with a bounded error while the bounds can be made arbitrarily small by adjusting the design variables. Simulation examples verified the effectiveness of the proposed approach.

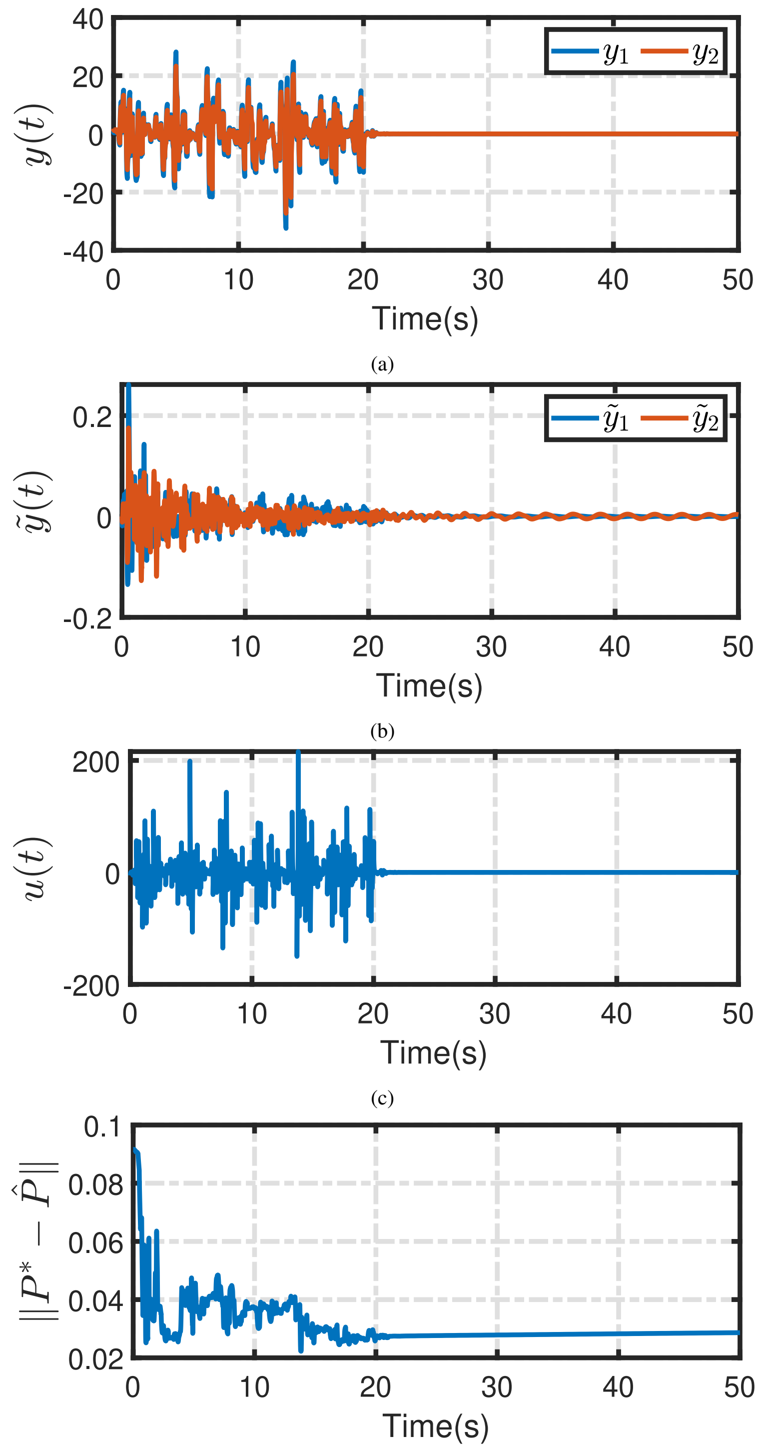
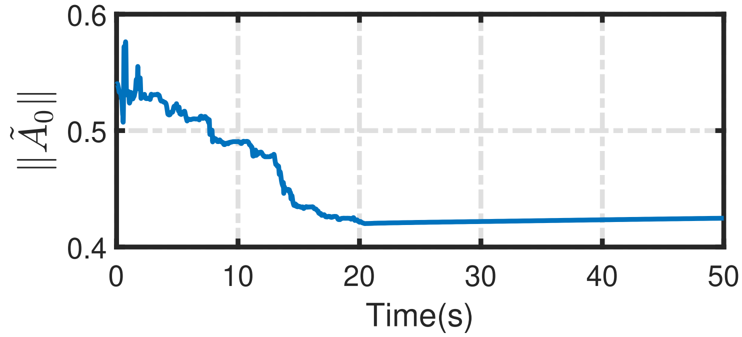
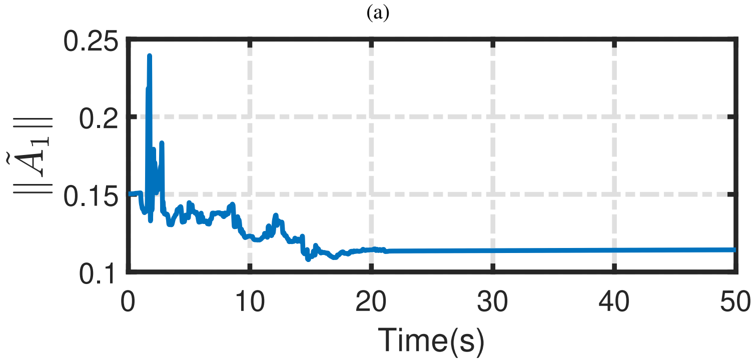


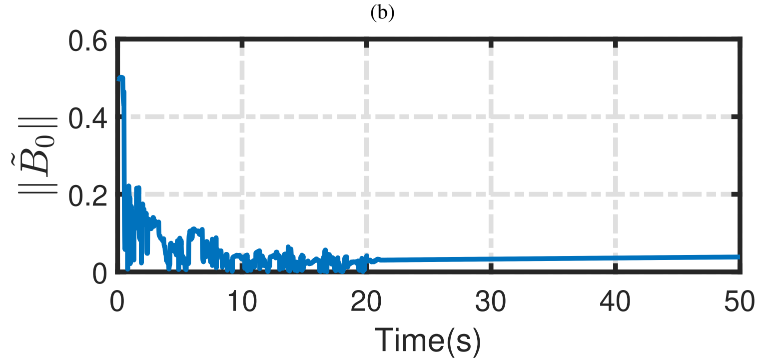
Figure 3. Performance of the optimal adaptive regulator. (a) System output. (b) Output estimation error. (c) Control input. (d) DARE solution error.



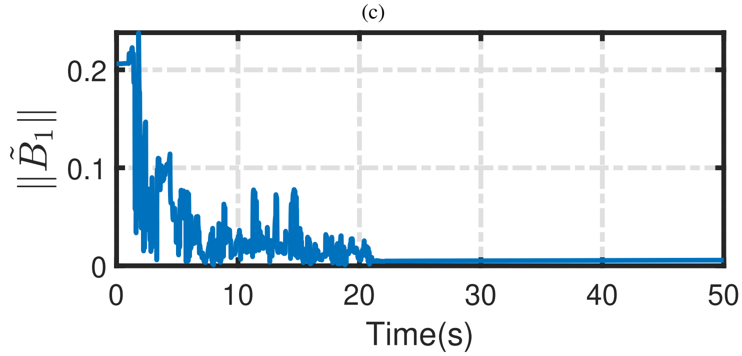
(a)



(b)



(c)



(d)

Figure 4. Norm of the parameter estimation errors. (a) $\|\tilde{A}_0\|$. (b) $\|\tilde{A}_1\|$. (c) $\|\tilde{B}_0\|$. (d) $\|\tilde{B}_1\|$.

REFERENCES

[1] Emilia Fridman. Introduction to Time-delay Systems: Analysis and Control. Springer, 2014.

[2] Y. Zhu, M. Krstic, and H. Su. Adaptive output feedback control for uncertain linear time-delay systems. IEEE Transactions on Automatic Control, 62(2):545–560, Feb 2017.

[3] Z. Artstein. Linear systems with delayed controls: A reduction. IEEE Transactions on Automatic Control, 27(4):869–879, August 1982.

[4] X. Cai, N. Bekiaris-Liberis, and M. Krstic. Input-to-state stability and inverse optimality of linear time-varying-delay predictor feedbacks. IEEE Transactions on Automatic Control, 63(1):233–240, Jan 2018.

[5] V. L. Kharitonov, S. . Niculescu, J. Moreno, and W. Michiels. Static output feedback stabilization: necessary conditions for multiple delay controllers. IEEE Transactions on Automatic Control, 50(1):82–86, Jan 2005.

[6] D. Ross and I. Flügge-Lotz. An optimal control problem for systems with differential-difference equation dynamics. SIAM Journal on Control, 7(4):609–623, 1969.

[7] O. Santos, S. Mondié, and V.L. Kharitonov. Linear quadratic suboptimal control for time delays systems. International Journal of Control, 82(1):147–154, 2009.

[8] Miroslav Krstic. Lyapunov tools for predictor feedbacks for delay systems: Inverse optimality and robustness to delay mismatch. Automatica, 44(11):2930 – 2935, 2008.

[9] Michael Basin and Jesus Rodriguez-Gonzalez. A closed-form optimal control for linear systems with equal state and input delays. Automatica, 41(5):915 – 920, 2005.

[10] Y. Fiagbedzi and A. Pearson. Feedback stabilization of linear autonomous time lag systems. IEEE Transactions on Automatic Control, 31(9):847–855, Sep. 1986.

[11] Shoulin Hao, Tao Liu, and Bin Zhou. Output feedback anti-disturbance control of input-delayed systems with time-varying uncertainties. Automatica, 104:8 – 16, 2019.

[12] Filippo Cacace and Alfredo Germani. Output feedback control of linear systems with input, state and output delays by chains of predictors. Automatica, 85:455 – 461, 2017.

[13] G. Tadmor. The standard H_∞ problem in systems with a single input delay. IEEE Transactions on Automatic Control, 45(3):382–397, 2000.

[14] S. Young Yoon and Z. Lin. Robust output regulation of linear time-delay systems: A state predictor approach. International Journal of Robust and Nonlinear Control, 26(8):1686–1704, 2016.

[15] V. Narayanan and S. Jagannathan. Event-triggered distributed control of nonlinear interconnected systems using online reinforcement learning with exploration. IEEE Transactions on Cybernetics, 48(9):2510–2519, 2018.

[16] Petros A Ioannou and Jing Sun. Robust adaptive control. Courier Corporation, 2012.

[17] W. Gao and Z. Jiang. Adaptive optimal output regulation of time-delay systems via measurement feedback. IEEE Transactions on Neural Networks and Learning Systems, 30(3):938–945, 2019.

[18] Kolmanovsky I. V. Orlov Y. and O. Gomez. Adaptive identification of linear time-delay systems: From theory toward application to engine transient fuel identification. International Journal of Adaptive Control and Signal Processing, 23(2):150–165, 2009.

[19] Romeo Ortega and R. Lozano. Globally stable adaptive controller for systems with delay. International Journal of Control, 47(1):17–23, 1988.

[20] Silviu-Iulian Niculescu and Anuradha M. Annaswamy. An adaptive smith-controller for time-delay systems with relative degree $n^* \leqslant 2$. Systems and Control Letters, 49(5):347 – 358, 2003.

[21] Anuradha Annaswamy Yildiray Yildiz, Ilya V. Kolmanovsky, and Diana Yanakiev. Adaptive posicast controller for time-delay systems with relative degree n⩾2. Automatica, 46(2):279 – 289, 2010.

[22] Nikolaos Bekiaris-Liberis, Mrdjan Jankovic, and Miroslav Krstic. Adaptive stabilization of lti systems with distributed input delay. International Journal of Adaptive Control and Signal Processing, 27(1-2):46–65, 2013.

[23] Changyun Wen Jing Zhou and Wei Wang. Adaptive backstepping control of uncertain systems with unknown input time-delay. Automatica, 45(6):1415 – 1422, 2009.

[24] Emilia Fridman Anton Selivanov and Alexander Fradkov. Passification-based adaptive control: Uncertain input and output delays. Automatica, 54:107 – 113, 2015.

[25] Kim-Doang Nguyen. An adaptive controller for uncertain nonlinear systems with multiple time delays. Automatica, 117:108976, 2020.

[26] Jonathan Chauvin Delphine Bresch-Pietri and Nicolas Petit. Adaptive control scheme for uncertain time-delay systems. Automatica, 48(8):1536 – 1552, 2012.

[27] Miroslav Krstic Yang Zhu and Hongye Su. Adaptive global stabilization of uncertain multi-input linear time-delay systems by pde full-state feedback. Automatica, 96:270 – 279, 2018.

[28] G. Strang. Introduction to linear algebra. Cambridge Press, 2016.

[29] Frank L Lewis, Draguna Vrabie, and Vassilis L Syrmos. Optimal Control. John Wiley & Sons, 2012.

[30] R. Moghadam and S. Jagannathan. Approximate optimal adaptive control of partially unknown linear continuous-time systems with state delay. In 2019 IEEE 58th Conference on Decision and Control (CDC), pages 1985–1990, 2019.

[31] De Schutter B. Busoniu L., Babuska R. and Ernst D. Reinforcement learning and dynamic programming using function approximators. CRC press, 2017.

[32] Kumpati S Narendra and Anuradha M Annaswamy. Stable adaptive systems. Courier Corporation, 2012.

[33] H. 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, 48(6):1017 – 1030, 2012.

[34] N. Cho, H. Shin, Y. Kim, and A. Tsourdos. Composite model reference adaptive control with parameter convergence under finite excitation. IEEE Transactions on Automatic Control, 63(3):811–818, 2018.

III. EVENT-TRIGGERED OPTIMAL ADAPTIVE CONTROL OF PARTIALLY UNKNOWN LINEAR CONTINUOUS-TIME SYSTEMS WITH STATE DELAY

Rohollah Moghadam, S.Jagannathan, and Vignesh Narayanan

Department of Electrical and Computer Engineering

Missouri University of Science and Technology

Rolla, Missouri 65409–0050

Email: moghadamr@umsystem.edu

ABSTRACT

The event-triggered optimal adaptive output feedback control of linear time-invariant systems with state delay is addressed by utilizing integral reinforcement learning (IRL) and integral temporal difference error (ITDE). The optimal control policy design and event-triggering threshold selection are treated as players in the game theoretic formulation. An infinite time horizon quadratic cost function is defined and a value functional, which includes the delayed output, is considered. By using the value functional and applying stationarity conditions using Hamiltonian, an output game delay algebraic Riccati equation (OGDARE) and optimal control policy using output are derived when the dynamics are known. Subsequently, when drift dynamics are considered uncertain, a hybrid learning scheme using measured output is proposed for tuning the value function parameters, which in turn is utilized to compute an estimated optimal control policy. The overall closed-loop system is shown to be asymptotically stable by selecting an event-triggering condition when system dynamics are both known and partially uncertain. A simulation example is provided to substantiate the efficacy of the theoretical claims.

1. INTRODUCTION

The study of time-delay systems has gained a significant interest because of its importance in practical applications such as network control systems (NCS) [1]. It is known that a delay in the feedback loop can affect the stability and performance of a physical system. Therefore, the delay differential equations and the control of time-delay systems has been extensively studied in the literature [2–17]. Designing a controller that is optimal and robust against the system uncertainties for time-delay systems is preferred. In addition, having the the system state vector available continuously, implies additional communication cost for the system when compared to output measurements. Therefore, the optimal control of time-delay systems with uncertain dynamics and with measured output is of utmost importance.

The optimal control tracking and regulation problems of time-delay systems have been reported in the literature [12, 18–20] when the system dynamics are known. An optimal controller is proposed in [19] for linear time-varying systems under multiple input delays by minimizing a finite horizon quadratic cost function. In [20], an optimal price-based decision making problem is addressed for a single-loop stochastic linear NCS in the presence of the network delay. These approaches [18–21] require system dynamics for the controller design which is a bottleneck in many industrial systems.

To overcome the need for accurate knowledge of the system dynamics, a reinforcement learning (RL) scheme [22] can be utilized for obtaining optimal control policy for such time-delayed systems by solving the fixed-point Bellman or Hamilton-Jacobi-Bellman (HJB) equations. In particular, the IRL method [23] has been developed to attain optimal control of uncertain discrete and continuous-time systems in the absence of delays [24]. Given a cost function, the IRL utilizes the Bellman equation to obtain the optimal control policy online. Despite relaxing the system dynamics, the traditional IRL requires continuous availability of the state vector.

The event-triggered control techniques [25–30], on the other hand, have been introduced to reduce the communication cost, and to some extent computational cost, in an NCS even when system has uncertain dynamics. In the event-triggered approach proposed in [25], the controller and the actuator use aperiodic mechanism and a zero-order-hold (ZOH) for the transmission of feedback information and control policy. An event-triggering condition is derived through the Lyapunov analysis and, if violated, an event is initiated for transmission of state and control inputs. The event sampled approach is similar to an RL agent interacting with the environment in aperiodic manner.

To reduce communication and computation cost and to attain optimality, the event-triggered optimal adaptive state feedback scheme with hybrid learning feature is introduced in [28, 31] for nonlinear continuous-time systems without delay. Note that the iterative optimal adaptive control (OAC) techniques [32] require an infinite number of iterations within a sampling interval to converge to an optimal policy, while the hybrid approach using state feedback converges to the optimal value by embedding a finite number of iterations within a sampling interval. In addition, the number of iterations changes with the duration of the sampling interval. In [33] an event-trigger output feedback controller is designed for robot manipulators with nonlinear dynamics. However, in all approaches, time delay in state vector is not considered.

In contrast, for time-delay systems due to the infinite dimensional nature, optimal policy depends upon both delay and the history of the system state vector. It has been shown in [3] that state and input delays can cause instability of the overall closed-loop systems unless the effect of the delays is considered explicitly. In [18, 34], an additional integral term based on the solution to the differential equation is needed for computing the control policy, which is not preferred for real-time control. Also, it causes pole-zero cancellation [35] when the system matrix is unstable. On the other hand, the finite dimensional, memoryless optimal control of time-delay systems with time-varying input delay is studied through traditional algebraic Ricatti equation (ARE) [21], provided a bound for the delay and control gain exists and complete knowledge of the system dynamics is known. However, state delay scenario is not considered. To the best of our knowledge, the problem of finding the OAC of linear time-invariant systems with state delay, free of integral term and in the presence of both known and partially uncertain dynamics, and aperiodic sampling of the system output has not been studied yet.

In this paper, an event-triggered finite dimensional and memoryless OAC of linear time-invariant (LTI) system with state delay by using output feedback is introduced. Two scenarios are considered: known system dynamics and partially unknown dynamics. First, under known system dynamics, the dynamic of the system is rewritten in the event-sampled two player zero-sum game formulation with control policy and control input error (or event trigger threshold) as two independent players. The aim is to minimize the control policy while maximizing the control input error or event sampled threshold [36]. The major benefit observed in the game theoretic formulation is the design of optimal policy and event-sampled threshold at the same time. Next, a quadratic cost function in terms of system output over an infinite time horizon is considered and a Lyapunov-Krasovskii function consisting of system output and its delayed value is introduced as the value functional. Then, Hamiltonian is formulated and, using the stationarity conditions, the Nash equilibrium to the game is determined in terms of the system output and solution to the output game delay algebraic Riccati equation (OGDARE). The OGDARE and event-trigger condition are derived to guarantee the stability of the closed-loop system using Luapunov analysis under the assumption that the system is controllable and observable.

Despite that the optimal control policy generated by our approach is free of any differential equation, solving OGDARE is involved and requires internal dynamics. In order to relax the need for partially unknown dynamics and the solution to OGDARE, the value functional is estimated by using both the IRL approach and output feedback. Integral temporal difference error (ITDE), which is derived from the difference between the actual and approximated value functional, is employed to find the tuning law for the unknown parameters of the value function, which includes the history of the delayed output. The event-trigger condition is obtained by using the game theoretic formulation and utilized in Lyapunov analysis.

Subsequently, a novel hybrid learning scheme is introduced by using measured output to tune the value function parameters during the inter-event sampled interval. The hybrid learning scheme updates the value function parameters at the sampled instances once and with a finite number of times, iteratively within the sampled interval. The number of iterations within the inter sampled interval is finite, but varies with its duration. Thus overall, the hybrid learning scheme relaxes the need for the significant number of iterations observed in traditional policy/value iteration techniques while still generating an optimal policy over time suitable for real-time control. The asymptotic stability of the overall closed-loop system is shown using Lyapunov theory. The Zeno-free behavior is also presented. The net result is the introduction of the event sampled for OAC of the time-delayed systems using output feedback.

The contributions of the paper are: 1) the development of a linear quadratic regulator using output feedback in the event sampled zero-sum formulation with control policy and control input error or threshold as the players, 2) introduction of a novel OGDARE, 3) development of an OAC via hybrid online learning feature using IRL in the event-sampled framework for the time-delay systems, and 4) The Lyapunov stability theory is utilized to analyze the closed-loop system, including Zeno-free behavior.

The paper is organized as follows. The background and problem formulation are presented in Section 2. The event-triggered linear quadratic regulator design for linear time-invariant systems with state delay when the dynamics are known is presented in Section 3 and, subsequently, Section 4 introduces event-triggered OAC using IRL and ITDE for tuning the parameters. In Section 5, simulation results is given and, conclusions is provided in Section 6.

Notation. Throughout the paper, I_n is the $n \times n$ identity matrix and the Euclidean vector norm denotes $\|.\|$ denotes. $A \otimes B$ is the Kronecker product of matrices A and B. The transpose of the matrix A is denoted as A^T. Maximum and minimum singular values of the matrix A are represented by $\delta_{max}(A)$ and $\delta_{min}(A)$, respectively. The minimum and maximum eigenvalue of the matrix A is defined as $\lambda_{min}(A)$ and $\lambda_{max}(A)$, respectively. The space of square integrable functions is denoted as L_2.

2. BACKGROUND AND PROBLEM FORMULATION

This section presents the optimal control of linear time-invariant system with state delay under the controllability and observability assumption. A cost function is defined as a quadratic function of the system output and control input. The optimal control policy is found for the case with available system dynamics, then, this assumption is relaxed by using an RL approach.

2.1. BACKGROUND

Let a linear time-invariant system with state delay (LTI-SSD) be defined as

$$\begin{cases} \dot{x}(t) = A_0 x\,(t) + Ax(t-d_x) + Bu(t) \\ y(t) = Cx(t),\;\; y(t-d_x) = Cx(t-d_x) \\ x(\theta) = \varphi(\theta) \qquad \theta \in [-d_x, 0] \end{cases}, \tag{1}$$

where $x(t) \in \mathbb{R}^n$, $u(t) \in \mathbb{R}^m$, and $y(t) \in \mathbb{R}^p$ denote the state, control input, and output, $A_0, A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$ and $C \in \mathbb{R}^{p\times n}$represent drift dynamics, input coefficient matrix, and output coefficient matrix, respectively, with $\varphi(.)$ being the initial function which is continuously differentiable and d_x is a known constant time delay.

Assumption 1. The LTI-SSD (1) is controllable and observable.

Remark 1. Note that Assumption 1 is required to ensure a solution to the delay algebraic Riccati equation, which will be given in the paper. The observability concept is that the initial value of the system state initial state can be derived through the measurement of the output of the system.

Remark 2. (Controllability and Observability Condition) ([3], Proposition 2.3, Theorem 2.4) The system (1) is controllable on $[0, t_1]$ for all $t_1 > nd_x$ if $rank(\bar{Q}) = n$, where

$$\begin{aligned}
&\bar{Q} = \left[\bar{Q}_1^1 \ldots \bar{Q}_1^n, \bar{Q}_2^2, \ldots, \bar{Q}_2^n, \ldots \bar{Q}_n^n\right] B \\
&\bar{Q}_1^1 = I, \bar{Q}_j^{k+1} = A_0^T \bar{Q}_j^k + A^T \bar{Q}_{j-1}^k, \\
&j = 1, \ldots, k+1, \quad k = 1, \ldots, n-1 \\
&\bar{Q}_j^k = 0, \quad j = 0 \quad or \quad j > k
\end{aligned} \tag{2}$$

Therefore, if the non-delay system, i.e., $\dot{x}(t) = A_0 x(t) + Bu(t)$ is controllable on $[t_0, t_1]$, or equivalently, the pair (A_0, B) is controllable, then, the delayed system (1) is controllable on $[t_0, t_1]$. Moreover, the system (1) is observable ([3], Preposition 2.4 , Theorem 2.6) on $[0, t_1]$ for all $t_1 > nd_x$ if $rank(\bar{P}) = n$, where

$$\begin{aligned}
&\bar{P} = \left[\bar{P}_1^1 \ldots \bar{P}_1^n, \bar{P}_2^2, \ldots, \bar{P}_2^n, \ldots \bar{P}_n^n\right] C^T \\
&P_1^1 = I, \bar{P}_j^{k+1} = A_0^T \bar{P}_j^k + A_1^T \bar{P}_{j-1}^k, \quad \begin{cases} j = 1, \ldots, k+1 \\ k = 1, \ldots, n-1 \end{cases} \\
&\bar{P}_j^k = 0, \quad j = 0 \quad or \quad j > k
\end{aligned} \tag{3}$$

Then, if the non-delay system of (1), i.e., $A = 0$ is observable on $[t_0, t_1]$, or equivalently, the pair (A_0, C) is observable, then (1) is observable on $[t_0, t_1]$.

2.2. PROBLEM FORMULATION

Consider the LTI-SSD (1). Assuming the output of the system is not continuously available, design a control policy $u(t) = -Ky(t_k)$, where t_k is the samples output and $K \in \mathbb{R}^{m\times p}$ is a user defined gain matrix which will be defined later, such that the closed-loop system is asymptotically stable, i.e., $\lim_{t\to\infty} x(t) \to 0$, while minimizing the quadratic cost function, $V(x)$, over infinite horizon given by

$$V(x(t)) = \int_t^{\infty} \left(y(\tau)^T \tilde{Q} y(\tau) + u(\tau)^T R u(\tau)\right) d\tau, \tag{4}$$

where $\tilde{Q}$ and $R \in \mathbb{R}^{m\times m}$ represent user defined positive definite matrices. The following Hamiltonian function is defined for (4) as

$$H(x, V, u) \equiv \frac{\partial V}{\partial t} + y^T \tilde{Q} y + u^T R u = 0, \tag{5}$$

where $\partial V/\partial t$ denotes the time derivative of the value function along the time. In [18, 34] the Hamiltonian equation (5) is called the Bellman type equation (BTE) for time-delay systems. Next, event-trigger mechanism utilized in the paper is introduced.

To develop the event-trigger mechanism, the Zero-order-hold (ZOH)-based approach is used to eliminate the need of continuous availability for the system output vector and, in the next section, the triggering threshold is derived using Lyapunov analysis.

As shown in Figure 1, in the actuator, a ZOH mechanism holds the control input during inter-event times until new output information is received. This makes the control input as a piece-wise continuous signal. Let the sampled state, from the event-triggered mechanism, to be sent to the controller at the sampling instant t_k be considered as $x(t_k)$. Then, the state measurement error can be illustrated by

$$e_x(t) = y(t_k) - y(t),\ t_k \leqslant t < t_{k+1} \tag{6}$$

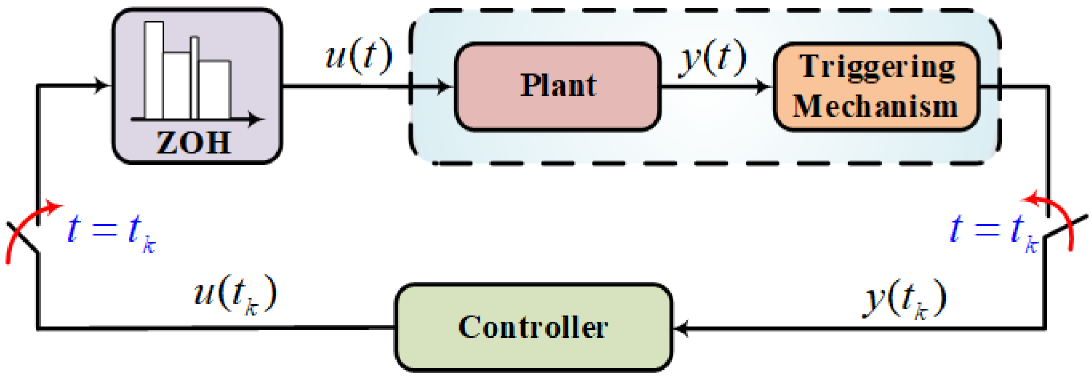


Figure 1. The ZOH-based event-triggered controller scheme.

Using the samples state $y(t_k)$, the sampled control input becomes $u(t_k) = Ky(t_k)$, where K is the controller gain and will be defined later. This sampled control input will be held at the actuator by ZOH mechanism to make the controller piece-wise continuous and it will get updated upon the arrival of the next sampled state. Now, using the sampled input $u(t_k)$, the LTI-SSD (1) can be expressed as

$$\dot{x}(t) = A_0x(t) + Ax(t - d_x) + Bu(t_k). \tag{7}$$

To proceed, let the error between the sampled control input, i.e., $u(t_k) = KCx(t_k)$ and actual control input, i.e., $u(t) = KCx(t)$ be defined as the control input error as

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

which from (6) one can conclude that $e_u(t) = Ke_x(t)$. Then, substituting (8) in (7), one gets

$$\dot{x}(t) = A_0x(t) + Ax(t - d_x) + Bu(t) + Be_u(t). \tag{9}$$

It can be seen from (9) that control input error, resulting from the aperiodic execution of the controller, acts as an external input to the system.

To obtain an optimal control input and a maximum threshold for $e_u(t)$, we consider the following dynamical system with a fictitious and continuous input $\hat{e}_u \in L_2$, given by

$$\dot{x}(t) = A_0 x(t) + Ax(t - d_x) + Bu(t) + B\hat{e}_u(t), \tag{10}$$

where the initial conditions for (10) are identical to (1). We will be treat the fictitious input $\hat{e}_u$ as a finite energy signal modeling the effect of control input error $e_u(t)$. Then, to incorporate the effect of the control input error, the cost function in (4) is redefined as

$$V = \int_t^\infty \left(y(\tau)^T \tilde{Q} y(\tau) + u(\tau)^T R u(\tau) - \gamma^2 \hat{e}_u^T \hat{e}_u \right) d\tau, \tag{11}$$

where γ denotes the user-defined attenuation constant. One can see that designing an optimal control input u while finding $\hat{e}_u(t)$ that maximizes the cost (11) results in optimizing the performance for the system (10). This problem formulation turns into the designing of an H_∞ controller for (10). Using the cost function (11), (5) turns into

$$\min_u \max_{e_u} \left\{ \frac{\partial V}{\partial t} + y^T \tilde{Q} y + u^T R u - \gamma^2 \hat{e}_u^T \hat{e}_u \right\} = 0. \tag{12}$$

We will utilize the solution of this optimization problem for the design of optimal control for the system (1), and incorporate the maximizer $\hat{e}_u$ as a threshold to bound the external input $e_u(t)$ in the event-triggered control framework for (1). Consequently, we design the triggering instants such that the event-triggered control system (9) mimics the optimal performance corresponding to the system (10).

Remark 3. Note that the control input error and the control input are considered as two independent players in the zero-sum game formulation (11). Although it appears that the two players are dependent, the worst case value of the control input error $\hat{e}_u$ is treated as a threshold for the triggering condition, which is viewed as an independent variable from the control input.

3. EVENT-TRIGGERED LINEAR QUADRATIC REGULATOR DESIGN

In this section, an optimal control approach is presented for LTI-SSD (1), using the output feedback. First, the system under consideration (1) is expressed as a two-player zero-sum game formulation in the event-sampled framework. Next, a Lyapunov-Krasovskii function consisting of system output and its delayed value is introduced as the value functional. Then, using the value functional, the BTE is derived. Using the BTE, the saddle-point solution to the game is obtained via stationarity conditions provided a solution to the OGDARE exists. Given a positive definite solution to OGDARE, using the Lyapunov analysis, it is demonstrated that the closed-loop system is asymptotically stable. The following lemma is used in the analysis.

Lemma 1. [37] Let M be a positive semi-definite square matrix and N be a positive definite square matrix, i.e., $v^T M v \geqslant 0$ and $v^T N v > 0$ for any vector $v \neq 0$. Then, $M + N$ is a positive definite square matrix.

To proceed, let the following Lyapunov-Krasovskii function be considered as the value function $V(x(t))$ defined in (4) as

$$V(x(t)) = y(t)^T P y(t) + \alpha \int_{t-d_x}^{t} y^T(s) y(s) ds \,, \tag{13}$$

where $P \in \mathbb{R}^{p\times p}$ is a symmetric positive definite matrix and $\alpha > 0$ is a positive scalar. Now, taking the first derivative of the value function (13) along with the system dynamic (10) and using Leibniz integral formula, the BTE (12) gives

$$\begin{aligned} BTE = {} & (A_0 x + A x(t-d_x) + Bu + B\hat{e}_u)^T C^T P y \\ & + y^T P C (A_0 x + A x(t-d_x) + Bu + B\hat{e}_u) \\ & + \alpha y^T y - \alpha y^T (t-d_x) y(t-d_x) \\ & + y^T \tilde{Q} y + u^T R u - \gamma^2 \hat{e}_u^T \hat{e}_u \end{aligned} \,. \tag{14}$$

To find the saddle point solution, the stationarity condition [38] is utilized for the above equation as $\partial(BTE)/\partial u = 0$ and $\partial(BTE)/\partial e_u = 0$, which yields

$$\begin{cases} u^*(t) = -R^{-1}B^T C^T P y(t) \\ e_u^*(t) = \frac{1}{\gamma^2} B^T C^T P y(t) \end{cases} . \tag{15}$$

where P is the solution to the following OGDARE given by

$$\begin{aligned} & A_0^T C^T PC + C^T PCA_0 + \tilde{Q} \\ & - C^T PC \left(B(2R^{-1} - \frac{2}{\gamma^2})B^T - \frac{1}{\alpha} AA^T \right) C^T PC = 0 \end{aligned} , \tag{16}$$

with $R > 0$ and $\tilde{Q} = \alpha C^T C + Q$, $Q \in \mathbb{R}^{n\times n}$ being a user defined positive definite matrix. By using Lemma 1, one can show that $\tilde{Q}$ is a positive definite matrix. The Riccati type equation (16) is termed as OGDARE because of the presence of the delayed state dynamics, i.e., A, in the equation. From (16) one can see that the derived Riccati type equation for the time delay system (1) is similar to game algebraic Riccati equation (GARE) for the non-delay systems and in the presence of disturbances. In this paper, the triggering instants and the control input are derived from the saddle point solution (u^*, e_u^*) given in terms of measured output.

Remark 4. Note that the proposed optimal controller of LTI-SSD is different from the one studied extensively in the literature [18, 34] as value functionals selected are not the same. The obtained OGDARE (16) does not depend upon the state delay, whereas it requires the coefficient matrix of the delayed state and, therefore, it is different from a traditional GARE of non-delay systems. Also, the control policy does not require an integral term or the solution to a differential equation. As discussed in [35], finding the integral term as the solution to a differential equation must be discarded as it involves pole-zero cancellation when the internal dynamics, A, is unstable. Moreover, numerical calculation of the integral term may cause instability of the overall closed-loop system.

In the following lemma, the Lyapunov analysis is employed to show that the the saddle point solution (15) makes the closed-loop system asymptotically stable if there exists a positive definite solution to the OGDARE (16).

Lemma 2. Consider the LTI-SSD (10) and under the aperiodic control input $u(t_k)$. Let Assumptions 1 be satisfied and there exists a positive definite matrix P as the solution of the OGDARE (16). Then, the control policy (15) renders the closed-loop system asymptotically stable provided $\lambda_{min}(C^T C - I_n) > 0$.

Proof. Consider the value function (13) as a Lyapunov-Krasovskii candidate. Then, the derivative of V along system trajectory and using the Leibniz integral formula gives

$$\begin{aligned} \dot{V}(x(t)) &= \dot{x}^T C^T PCx + x^T C^T PC\dot{x} \\ &\quad + \alpha y^T(t)y(t) - \alpha y^T(t-d_x)y(t-d_x). \end{aligned} \tag{17}$$

Substituting the system dynamics (10) and saddle point solution (15) into (17) gives

$$\begin{aligned} \dot{V} = &(A_0 x(t) + Ax(t-d_x) - BR^{-1}B^T C^T PCx(t) \\ &+ \frac{1}{\gamma^2} BB^T C^T PCx(t))^T C^T PCx \\ &+ x^T C^T PC(A_0 x(t) + Ax(t-d_x) \\ &- BR^{-1}B^T C^T PCx(t) + \frac{1}{\gamma^2} PBB^T C^T PCx(t)) \\ &+ \alpha x^T(t) C^T Cx(t) - \alpha x^T(t-d_x) C^T Cx(t-d_x) \end{aligned} \tag{18}$$

which yields

$$\begin{aligned} \dot{V} = &x^T(A_0^T C^T PC + C^T PCA_0 + \alpha C^T C \\ &-2C^T PCBR^{-1}B^T C^T PC + \frac{2}{\gamma^2} C^T PCBB^T C^T PC)x \quad . \\ &- \alpha x^T(t-d_x) C^T Cx(t-d_x) + 2x(t-d_x)^T A^T C^T PCx \end{aligned} \tag{19}$$

Adding and subtracting $\frac{1}{\alpha}C^TPCAA^TC^TPC$ and $\alpha x(t-d_x)^Tx(t-d_x)$ and after simplification, (19) can be written as

$$\begin{aligned}\dot{V} &= x^T(A_0^TC^TPC + C^TPCA_0 + \alpha C^TC - 2C^TPCBR^{-1}B^TC^TPC \\ &+\frac{2}{\gamma^2}C^TPCBB^TC^TPC + \frac{1}{\alpha}C^TPCAA^TC^TPC\Big)x \\ &- \alpha x^T(t-d_x)(C^TC - I_n)x(t-d_x) \\ &- \frac{1}{\alpha}\left[A^TC^TPCx - \alpha x(t-d_x)\right]^T\left[A^TC^TPCx - \alpha x(t-d_x)\right]\end{aligned}, \tag{20}$$

which by substituting the OGDARE (16) and since the last term of (20) is always negative semi-definite, (20) can be expressed by

$$\dot{V} \leqslant -x^TQx - \alpha x^T(t-d_x)(C^TC - I_n)x(t-d_x)\ . \tag{21}$$

This states that the derivative of the Lyapunov-Krasovskii function is negative definite, i.e., $\dot{V} < 0$ provided $\lambda_{min}(C^TC - I_n) > 0$. This results in asymptotic stability of the closed-loop system. ■

Remark 5. As shown in Lemma 2, the system is asymptotically stable and, therefore, as shown in [39] for non-delay systems and in [40] for time-delay systems, it becomes an input-to-state stable system.

In the following theorem, the triggering condition is defined by using Lyapunov analysis.

Theorem 1. Consider the LTI-SSD (1) under event-sampled framework along with the cost function (11). Suppose there exists a positive definite solution to OGDARE (16). Let the triggering condition be given by

$$\|e_u(t)\| \leqslant \|e_u^*(t)\|, \quad t \in [t_k, t_{k+1}), \quad \forall k \in 1, 2, \ldots, \tag{22}$$

where e_u^* is defined in (15). Then, the closed-loop system event-triggered time-delay system becomes asymptotically stable.

Proof. Let the value function (13) be defined as a Lyapunov-Krasovskii candidate. Taking the first derivative of V along system trajectory (6) and substituting the control input from (15) gives

$$\begin{aligned}\dot{V} &= (A_0x(t) + Ax(t-d_x) - BR^{-1}B^TPCx(t) \\ &+ Be_u(t))^TC^TPCx + x^TC^TPC(A_0x(t) \\ &+ Ax(t-d_x) - BR^{-1}B^TPCx(t) + Be_u(t)) \\ &+ \alpha x^T(t)C^TCx(t) - \alpha x^T(t-d_x)C^TCx(t-d_x)\end{aligned}, \tag{23}$$

which by adding and subtracting $\frac{1}{\alpha}C^TPCAA^TC^TPC$ and $\alpha x(t-d_x)^Tx(t-d_x)$ and after simplification, (23) turns into

$$\begin{aligned}\dot{V} &= x^T\left(A_0^TC^TPC + C^TPCA_0 - 2C^TPCBR^{-1}B^TC^TPC\right. \\ &\left.+\alpha C^TC + \frac{1}{\alpha}C^TPCAA^TC^TPC\right)x + 2x^TC^TPCBe_u(t) \\ &-\alpha x^T(t-d_x)(C^TC - I_n)x(t-d_x) \\ &-\frac{1}{\alpha}\left[A^TC^TPCx - \alpha x(t-d_x)\right]^T\left[A^TC^TPCx - \alpha x(t-d_x)\right]\end{aligned}. \tag{24}$$

Now, using the OGDARE (16), (24) becomes

$$\dot{V} \leqslant -x^T\left(Q + \frac{2}{\gamma^2}C^TPCBB^TC^TPC\right)x + \|2x^TC^TPCBe_u(t)\| \tag{25}$$

which can be simplified as

$$\dot{V} \leqslant -\varphi_1(\|y\|) + \varphi_2(\|y\|)\|e_u(t)\| \tag{26}$$

where $\varphi_1(\|y\|) = \frac{\|y\|^2}{\|C\|^2}\|Q + \frac{2}{\gamma^2}C^TPCBB^TC^TPC\|$ and $\varphi_2(\|y\|) = \|y\|\|2PCB\|$. From (26), $\dot{V} < 0$ if $\|y\| > \varphi_1^{-1}(\varphi_2(\|y\|))\|e_u\|$. This implies that V is an ISS Lyapunov function [41].

Using the triggering condition (22), one has

$$\dot{V} \leqslant -\varphi_1(\|y\|) + \varphi_2(\|y\|)\|e_u^*(t)\| \tag{27}$$

Now, using the definition of $e_u^*(t)$ in (15) and the definition of $\varphi_2(\|x\|)$, equation (27) yields

$$\dot{V} \leqslant -\left(\varphi_1(\|y\|) - \frac{1}{2\gamma^2}(\varphi_2(\|y\|))^2\right) \tag{28}$$

Therefore, one can design the parameter γ such that $\varphi_1(\|y\|) > \frac{1}{2\gamma^2}(\varphi_2(\|y\|))^2$. This makes the first derivative of the Lyapunov-Krakovskii function negative, i.e. $\dot{V} < 0$, which implies the asymptotic stability of the closed-loop system employing the triggering condition (22). Since $y = Cx$, this implies that the system state vector tends to be zero asymptotically. ■

The following theorem shows that the event-trigger condition, derived from the zero-sum game, is Zeno-free.

Theorem 2. Consider the LTI-SSD (1) with the cost function (4). Let P be a symmetric positive definite solution to the OGDARE (16) and the triggering condition (22) be employed to generate the event sampled control input. Then, the event-triggering condition (22) is Zeno-free.

Proof. To show a positive inter-event time, consider any two trigger instants as $\{t_k, t_{k+1}\}$. Then, using the measurement error (6), $\|e_u\| > 0$ for all $t \subseteq (t_k, t_{k+1})$ and the inequatily $\frac{d}{dt}\|e_u\| = \frac{d}{dt}\left(e_u^Te_u\right)^{1/2} = \frac{e_u^T\dot{e}_u}{\|e_u\|} \leqslant \frac{\|e_u\|\|\dot{e}_u\|}{\|e_u\|} = \|\dot{e}_u\|$ holds. Note that on the time interval $t \in [t_k, t_{k+1})$, one has

$$\begin{cases} \dot{e}_u = K\dot{x} \\ e_u(t_k) = 0 \end{cases}, \tag{29}$$

where K is a constant control gain defined by (15) and given by $K = -R^{-1}B^TC^TPC$. Then, using the system dynamics (1), $\forall t \in (t_k, t_{k+1})$, one can observe that $\|\dot{e}_u(t)\| \leqslant \|A_0\| \|e_u(t)\| + \zeta_k$ where $\zeta_k = max_{t\in(t_k,t_{k+1})}\|KA_0x(t_k) + KAx(t_k - d_x) + KBKx(t_k)\|$. Therefore, using $\frac{d}{dt}\|e_u\| \leqslant \|\dot{e}_u\|$, one has

$$\|e_u\| \leqslant \frac{\zeta_k}{\|A_0\|}\left(e^{\|A_0\|(t-t_k)} - 1\right) \tag{30}$$

At $t = t_{k+1}$, one has $\frac{1}{\gamma^2}\|B^TC^TPCx\| \leqslant \|e_u\| \leqslant \frac{\zeta_k}{\|A_0\|}\left(e^{\|A_0\|(\tau_k)} - 1\right)$ which results in

$$\tau_k \geqslant \frac{1}{\|A_0\|}\ln\left(\frac{\|A_0\|}{\gamma^2\zeta_k}\|B^TC^TPCx\| + 1\right) \tag{31}$$

Therefore, from the expression of the inter-event time (31), it can be concluded that while estimated parameters affect the time between consecutive event-triggering instants, the inter-event time is positive and prevents continuous triggering. ■

Remark 6. According to Lemma 2, the time-delay system (10) by considering the saddle point solution (15) is asymptotically stable. Now, to show that the time delay system (1) is input to state stable, one can show that using the OGDARE (16) and the value function (13) as the Lyapunov candidate function, the first derivative of the Lyapunov function along the system dynamics (1) becomes $\dot{V} \leqslant -x^T(Q + \frac{2}{\gamma^2}C^TPCBB^TC^TPC)x - \alpha x^T(t-d_x)(C^TC - I_n)x(t-d_x)$. This implies that, when the conditions provided in Lemma 2 are satisfied, the time-delay system (1) with the optimal control u^*, given in (15), is asymptotically stable provided $e_u = 0$. Therefore, as shown in Proposition 2.5 [40], the time delay system (1) is input to state stable. Consequently, asymptotic stability of the system during the inter-event times given in Theorem 1 results in the asymptotic stability of the system in the overall time intervals.

It is shown in Theorem 1 that, using the triggering condition, the closed-loop system is asymptotically stable when the dynamics of the system are known. In the next section, the controller and triggering condition is designed such that the system dynamics are partially unknown while the coefficient matrix associated with the control input is known.

4. HYBRID EVENT-TRIGGERED OPTIMAL ADAPTIVE CONTROL MECHANISM

The IRL approach [23] is integrated with the hybrid learning and event-trigger mechanism to obtain the OAC of LTI-SSD (1) under partially unknown dynamics. Here, integral IRL is utilized in conjunction with approximate dynamic programming under a game theoretic framework to find the optimal control policy and event-trigger condition.

It is important to notice that traditional dynamic programming works in a backward and off-line manner when the system dynamics are known, whereas an IRL approach helps in obtaining approximate optimal solutions in a forward-in-time manner even when the dynamics are uncertain. In contrast to using policy or value iterative techniques to solve the Bellman or HJB equation, the TDE is an alternate approach of finding the optimal control policy online and forward in time. Actually, the TDE can be treated as a predictor error between observed and predicted performance in response to the control policy applied to the system. However, finding the solution to the Lyapunov equation arising from TDE makes it hard to solve.

To overcome this issue, one can estimate value function by using a parametric approximator whose parameters can be adjusted by using integral TDE (ITDE). Then, a PE condition is required to ensure the value function parameters converge asymptotically to the target optimal value. Instead, a hybrid-based parameter tuning law is proposed using the measured output which not only reduces the need for continuous measurement of the system information, but also converges to the target parameters faster [28], since a finite number of parameter updates are introduced within each sampling instant.

In this paper, the ITDE is utilized to obtain OAC of time-delay system (1) under partially unknown dynamics. The technique has two steps: critic and actor update. In IRL, the measured output vector at specific sampling times $t + i\Delta t$ with $i \in \mathbb{N}$ is used. The ITDE is employed to update the critic function with the current stabilizing controller. Then, the critic function is used to update the actor. In this procedure, only measured output is utilized for updating the critic parameters, and, no knowledge of the system internal dynamics is required. Here, the ITDE, which includes the history of the delayed output, is utilized to tune the value function parameters under a PE condition in the proposed event-sampled game theoretic framework.

Due to the varying inter-event interval , the number of iterations performed on the critic function parameters will change over time. As a consequence, the stability of the closed-loop system is a challenge unless suitable parameter update law is employed. Traditional adaptive control does not use a combination of time-driven and iterative updates. Here, an event-sampled adaptive approach is employed to find the estimated value function and control policy.

By using the IRL, the Bellman equation can be represented as

$$V(y(t)) = \int_{t}^{t+\Delta t} \left(y^T \tilde{Q} y + u^T R u - \gamma^2 e_u^T e_u \right) d\tau + V(y, t + \Delta t), \tag{32}$$

with Δt as a small and fixed time interval. Now, the value function (13) can be expressed as

$$V(y(t)) = \theta^T \sigma(y) + \alpha \int_{t-d_x}^{t} y^T(s) y(s) ds \tag{33}$$

with $\theta = vech(P)^T \in \mathbb{R}^{\frac{1}{2}n(n+1)}$, where $vech(P)$ implies the vector created from the matrix P, which is symmetric. The regression function in terms of the output vector is denoted as $\sigma(y) = y \otimes y$ where $\times$ is Kronecker product.

Since the system dynamics are partially unknown, the actual parameters of the optimal value function, i.e. V^*, cannot be determined. Instead, the value function can be estimated by using the measured output as

$$\hat{V}(y(t)) = \hat{\theta}^T \sigma(y) + \alpha \int_{t-d_x}^{t} y^T(s) y(s) ds\,, \tag{34}$$

where $\hat{\theta} \in \mathbb{R}^{\frac{1}{2}n\times(n+1)}$ denote the estimated parameters. It should be noted that the value function estimation has an integral term of delayed value of the output from the state delay.

Remark 7. The second term in the value function (13) is known and, therefore, is not estimated.

Next, using the estimated value function, the following ITDE due to the difference between the actual and estimated parameters can be defined as

$$e_{ITDE} = \hat{V}(y, t+\Delta t) - \hat{V}(y,t) + \int_{t}^{t+\Delta t} \left(y^T \tilde{Q} y + u^T R u - \gamma^2 e_u^T e_u\right) d\tau \tag{35}$$

which becomes

$$\begin{aligned} e_{ITDE} = & \int_{t}^{t+\Delta t} \left(y^T \tilde{Q} y + u^T R u - \gamma^2 e_u^T e_u\right) d\tau + \hat{\theta}^T \Delta\sigma(y) \\ & + \alpha \left(\int_{t+\Delta t-d_x}^{t+\Delta t} y^T(s) y(s) ds - \int_{t-d_x}^{t} y^T(s) y(s) ds \right) \end{aligned} \tag{36}$$

where $\Delta\sigma(y) = \sigma(y(t+\Delta t)) - \sigma(y(t))$. Now, an appropriate update law should be designed for $\hat{\theta}$ such that the ITDE (36) tends to zero, eventually. To this end, considering $E = \frac{1}{2}\left\|e_{ITDE}\right\|^2$ and using the gradient descend method, the following update law can be constructed at the sampled instants and inter-sampled interval as

$$\begin{cases} \hat{\theta}^+ = \hat{\theta}(t_k) - \beta \dfrac{\Delta\sigma(y)}{\left(1+\Delta\sigma(y)^T \Delta\sigma(y)\right)^2} e_{ITDE}^T(t_k),\ t = t_k \\ \dot{\hat{\theta}} = -\beta \dfrac{\Delta\sigma(y(t_k))}{\left(1+\Delta\sigma(y)^T \Delta\sigma(y)\right)^2} e_{ITDE}^T(t_k),\ t \in (t_k, t_{k+1}) \end{cases}, \tag{37}$$

where β is a scalar gain determining the convergence rate. The right derivative of the estimated parameter $\hat{\theta}$ at the event-triggering instant t_k is denoted as θ^+. Note that to ensure the convergence of the estimated parameters, the regression vector should be persistently exciting.

The overall event-triggered optimal adaptive controller with hybrid learning is shown in Figure 2. The value function estimator is located at the event-trigger mechanism as well as the controller in order to prevent transmission of parameters from the controller to the event-triggering mechanism. These two value function estimators are synchronized using the same initial conditions. Moreover, in the parameter tuning law for the hybrid scheme (37), the value of the regression vector $\Delta\sigma(y)$ and the ITDE, e_{ITDE}, at the event-sampled time instants t_k is utilized in the update law for the inter-sampled time interval $t \in [t_k, t_{k+1})$.

Now, updating the actor with the estimated critic, (15) can be written as

$$u(t) = -R^{-1}B^T C^T \nabla^T \sigma(y)\hat{\theta}(t) \tag{38}$$

where $\nabla\sigma$ denotes the derivative of the regression function σ with respect to the output y. The the worst case event-triggering threshold can be expressed as

$$\hat{e}_u(t) = \frac{1}{\gamma^2}B^T C^T \nabla^T \sigma(y)\hat{\theta}(t) \tag{39}$$

Defining $\tilde{\theta} = \theta - \hat{\theta}$ as the parameter estimation error and using (36) and (33), gives

$$e_{ITDE} = -\tilde{\theta}^T \Delta\sigma(y) \tag{40}$$

Then, the parameter estimation error dynamics becomes

$$\begin{cases} \tilde{\theta}^{+} = \tilde{\theta}(t_k) + \beta \dfrac{\Delta\sigma(y)}{(1+\Delta\sigma(y)^T\Delta\sigma(y))^2} e_{ITDE}^T(t_k) & t = t_k \\ \dot{\tilde{\theta}} = \beta \dfrac{\Delta\sigma(y)}{(1+\Delta\sigma(y)^T\Delta\sigma(y))^2} e_{ITDE}^T(t_k) & t \in [t_k, t_{k+1}) \end{cases} \tag{41}$$

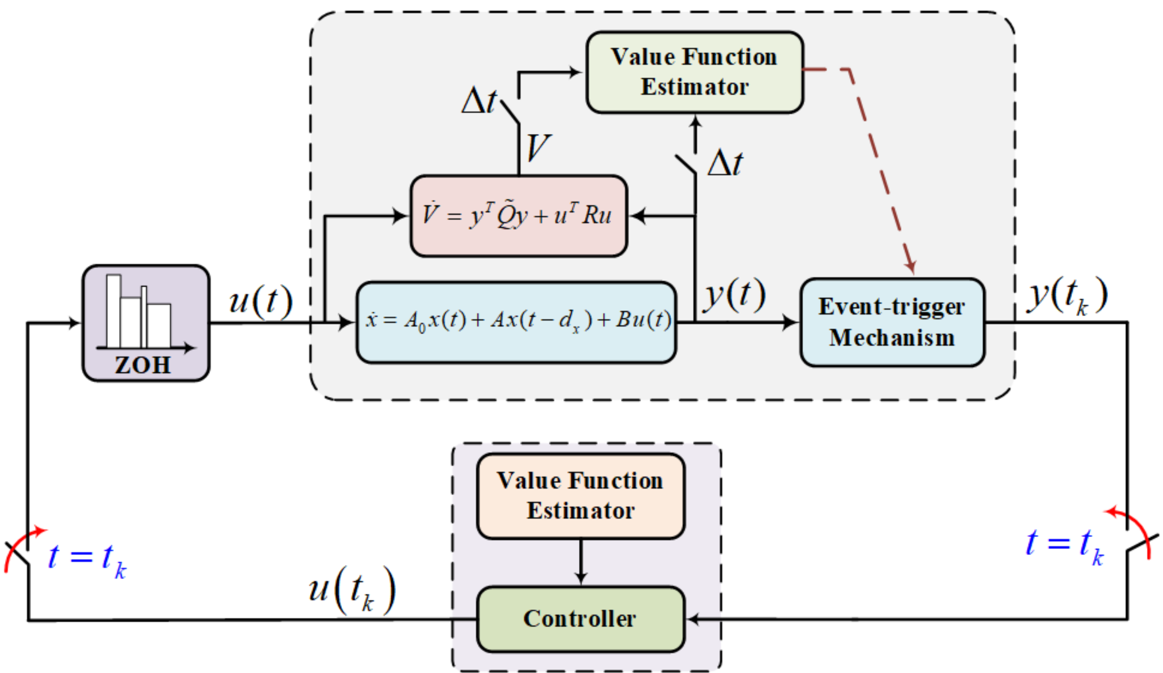


Figure 2. The hybrid event-triggered optimal adaptive controller scheme when internal dynamics are unknown.

In the following theorem, the stability of the overall closed-loop system by considering partially unknown system dynamics with aperiodic controller is presented.

Theorem 3. Consider the LTI-SSD (1) with the cost function defined in (4). Let the control input policy and the parameter update law be given by (38) and (37), respectively. The initial value function parameter vector, $\hat{\theta}(0)$, is defined in a compact set such that the initial control input be admissible. Let the triggering condition be given by

$$\|e_u(t)\| \leqslant \|\hat{e}_u(t)\| = \frac{1}{\gamma^2} B^T \nabla^T \sigma(y)\hat{\theta}, \quad t \in [t_k, t_{k+1}), \quad \forall k \in 1, 2, \ldots, \tag{42}$$

Then, the system state and parameter estimation errors are ultimately bounded provided $\beta < 2$ and $\frac{\beta}{\rho} > \|B\|^2\|\nabla^T\sigma(y(t_k))\|^2$ where $\frac{1}{\rho} = \inf_{\forall t\in(t_k,t_{k+1})}\left(\frac{\Delta\sigma(y)\Delta\sigma(y)^T}{\left(1+\Delta\sigma(x)^\top\Delta\sigma(y)\right)^2}\right)$. Moreover, when the regression vector is persistently exciting, the state of the system and the parameter estimation error converges to zero asymptotically as $k \to \infty$, where k represents the event number.

Proof. The proof has two parts as the update law is considered for two different time intervals, one during the inter-event times and another one at the triggering instants. First, it is shown that the system state vector and the parameter estimation error is bounded at the triggering instants. Then, the boundedness is shown during the triggering instants.

4.1. BOUNDEDNESS AT THE TRIGGERING INSTANTS, $T = T_K$

To proceed, the following Lyapunov candidate function is defined as

$$\mathcal{L} = V_1(y) + V_2(\tilde{\theta}) \tag{43}$$

where $V_1(y) = y(t)^TPy(t) + \alpha\int_{t-d_x}^{t} y^T(s)y(s)ds$ and $V_2(\tilde{\theta}) = \frac{1}{2}\tilde{\theta}^T\tilde{\theta}$.

For the proof of part 1, the first difference of (43) becomes

$$\Delta\mathcal{L} = \Delta V_1(y) + \Delta V_2(\tilde{\theta}) \tag{44}$$

The drivative of the first term can be written as $\Delta V_1(y) = V_1(y(t^+)) - V_1(y(t))$ which becomes $\Delta V_1(y) = y(t^+)^TPy(t^+) + \alpha\int_{t-d_x}^{t} y(t^+)^Ty(t^+)ds - y(t)^TPy(t) - \alpha\int_{t-d_x}^{t} y^T(s)y(s)ds$ and at the triggering instant $y(t^+) = y(t)$ which results in $\Delta V_1(y) = 0$. Then, for the second term of (44), one has

$$\Delta V_2(\tilde{\theta}) = \frac{1}{2}(\tilde{\theta}^+)^T\tilde{\theta}^+ - \frac{1}{2}\tilde{\theta}^T\tilde{\theta} \tag{45}$$

By the aid of (41) and (40), (45) results in

$$\begin{aligned}\Delta V_2(\tilde{\theta}) = &-\beta\tilde{\theta}^T\frac{\Delta\sigma(y)\Delta\sigma(y)^T}{\left(1+\Delta\sigma(y)^\top\Delta\sigma(y)\right)^2}\tilde{\theta}\\ &+\beta^2\tilde{\theta}^T\frac{\Delta\sigma(y)\Delta\sigma(y)^T\Delta\sigma(y)\Delta\sigma(y)^T}{2\left(1+\Delta\sigma(y)^\top\Delta\sigma(y)\right)^4}\tilde{\theta}\end{aligned}\tag{46}$$

which by taking the norm and using the inequality $\frac{\beta^2\|\Delta\sigma(y)^T\Delta\sigma(y)\|}{2\left(1+\Delta\sigma(y)^\top\Delta\sigma(y)\right)^4} \leqslant \frac{\beta^2}{2\left(1+\Delta\sigma(y)^\top\Delta\sigma(y)\right)^2}$, (46) can be upper bounded by

$$\begin{aligned}\Delta V_2(\tilde{\theta}) \leqslant &-\frac{\beta}{\left(1+\Delta\sigma(y)^\top\Delta\sigma(y)\right)^2}\|\tilde{\theta}^T\Delta\sigma(y)\|^2\\ &+\frac{\beta^2}{2\left(1+\Delta\sigma(y)^\top\Delta\sigma(y)\right)^2}\|\tilde{\theta}^T\Delta\sigma(y)\|^2\end{aligned}\tag{47}$$

Under the persistent excitation of the regression vector and considering $\frac{1}{\rho} = \inf_{\forall t\in(t_k,t_{k+1})}\left(\frac{\Delta\sigma(y)\Delta\sigma(y)^T}{\left(1+\Delta\sigma(x)^\top\Delta\sigma(y)\right)^2}\right)$, (47) turns into

$$\Delta V_2(\tilde{\theta}) \leqslant -\frac{\beta}{\rho}\left(1-\frac{\beta}{2}\right)\left\|\tilde{\theta}\right\|^2 \tag{48}$$

which implies $\Delta\mathcal{L} \leqslant 0$, if $\beta < 2$ at triggering instants. Since the first difference of the Lyapunov function does not include the output vector, it becomes negative semi-definite. However, since the output of the system does not change at the triggering instants and it will be proven that it becomes bounded during the inter-event times, one can conclude that the output vector becomes bounded at the triggering instants. And, since eventually the output vector during the inter-event time converges to zero, one can conclude that the overall derivative of the Lyapunov function becomes negative definite.

4.2. BOUNDEDNESS DURING THE INTER-EVENT TIMES, $T \in [T_K, T_{K+1})$

For the proof of part 2, the first derivative of the Lyapunov candidate function (43), during the inter-event times, becomes

$$\dot{\mathcal{L}} = \dot{V}_1(y) + \dot{V}_2(\tilde{\theta}) \tag{49}$$

The first derivative of $V_1(y)$, by considering the estimated control input and estimated control input error, can be written as

$$\begin{aligned}\dot{V} = &(A_0 x(t) + Ax(t-d_x) + B\hat{u} + B\hat{e}_u)^T C^T PCx \\ &+ x^T C^T PC(A_0 x(t) + Ax(t-d_x) + B\hat{u} + B\hat{e}_u) \ , \\ &+ \alpha x^T(t) C^T Cx(t) - \alpha x^T(t-d_x) C^T Cx(t-d_x)\end{aligned} \tag{50}$$

By adding and subtracting $2x^T C^T PCBu^*$ and $2x^T C^T PCBe_u^*$ and by using the procedure presented in the proof Lemma 2, (50) yields

$$\begin{aligned}\dot{V}_1 \leqslant &-x^T Qx + 2x^T C^T PCB(\hat{u} + \hat{e}) \\ &- 2x^T C^T PCB(u^* + e_u^*) - \alpha x^T(t-d_x)(C^T C - I_n)x(t-d_x)\end{aligned} \tag{51}$$

which, from the definition of the control input error in (8), can be written as

$$\begin{aligned}\dot{V}_1 \leqslant &-x^T Qx + 2x^T C^T PCB(\hat{u}(t_k) - u^*(t_k)) \\ &- \alpha x^T(t-d_x)(C^T C - I_n)x(t-d_x)\end{aligned} \tag{52}$$

Defining $\tilde{u}(t_k) = u^*(t_k) - \hat{u}(t_k) = -R^{-1}B^T \nabla^T \sigma(y(t_k))\tilde{\theta}(t_k)$ and using $-y^T PB = u^{*T} R$ from (15), (52) turns into

$$\dot{V}_1 \leqslant -x^T Qx - 2u^{*T} R\tilde{u}(t_k) - \alpha x^T(t-d_x)(C^T C - I_n)x(t-d_x) \tag{53}$$

Using the Young's inequality, (53) gives

$$\begin{aligned}\dot{V}_1 \leqslant & -x^T Qx + u^{*T} u^* \\ & + \tilde{\theta}^T \nabla\sigma(y(t_k)) BB^T \nabla^T \sigma(y(t_k)) \tilde{\theta} \\ & - \alpha x^T(t-d_x)(C^T C - I_n) x(t-d_x)\end{aligned} \tag{54}$$

Next, considering the first derivative of the second term in the Lyapunov candidate function (43) as $\dot{V}_2(\tilde{\theta}) = \tilde{\theta}^T \dot{\tilde{\theta}}$ and using the parameter estimation error (41) and (40), results in

$$\dot{V}_2 = -\beta \tilde{\theta}^T \frac{\Delta\sigma(y)\Delta\sigma(y)^T}{\left(1 + \Delta\sigma(y)^\top \Delta\sigma(y)\right)^2} \tilde{\theta} \tag{55}$$

Summing up (55) and (54), taking the norm and, after some simplification, one has

$$\begin{aligned}\dot{\mathcal{L}} \leqslant & -\delta_1(\|y\|^2) - \alpha\lambda_{min}(C^T C - I_n)\|x^T(t-d_x)\|^2 \\ & - \left(\frac{\beta}{\rho} - \|B\|^2 \|\nabla^T \sigma(y(t_k))\|^2\right) \|\tilde{\theta}\|^2\end{aligned} \tag{56}$$

where $\delta_1(\|y\|^2) = \frac{1}{\|C\|} x^T Qx - u^{*T} u^*$. This implies that if $\frac{\beta}{\rho} > \|B\|^2 \|\nabla^T \sigma(y(t_k))\|^2$ and $\lambda_{min}(C^T C - I_n) > 0$, then, considering the results for both cases of inter-event times and event-triggering instants yields $\dot{\mathcal{L}} < 0$, which guarantees the asymptotic stability of the overall closed loop system under the satisfaction of PE condition. Since the output goes to zero and using the output equation, $y = Cx$ with C a constant matrix, the system state vector converges to zero asymptotically. ■

Remark 8. Note that when the dynamics of the system are unknown, the same procedure provided in Theorem 2 can be done to prove that the event-triggering mechanism is Zeno-free and the minimum time interval is positive.

Remark 9. In a case where the system state vector is measurable, i.e., $C = I_n$, the same procedure can be followed by defining the cost function as

$$V(x(t)) = \int_t^\infty \left(x(\tau)^T \tilde{Q} x(\tau) + u(\tau)^T R u(\tau) \right) d\tau, \tag{57}$$

and the value function as

$$V(x(t)) = x(t)^T P x(t) + \alpha \int_{t-d_x}^t x^T(s) x(s) ds\,, \tag{58}$$

where P is the solution of the following game delay algebraic Riccati equation

$$A_0^T P + P A_0 + \tilde{Q} - P \left(2B(R^{-1} - \frac{1}{\gamma^2}) B^T - \frac{1}{\alpha} A A^T \right) P = 0\,, \tag{59}$$

Then, using the Bellman type equation for state feedback, the saddle point solution becomes

$$\begin{cases} u^*(t) = -R^{-1} B^T P x(t) \\ e_u^*(t) = \frac{1}{\gamma^2} B^T P x(t) \end{cases} . \tag{60}$$

Similarly, for the case when the internal dynamics of the system are uncertain, steps given in this section can be repeated by assuming $C = I_n$.

Next, it is shown that the estimated control input converges to the optimal one when PE condition is satisfied and parameter estimation error converges to zero asymptotically.

Theorem 4. Consider the LTI-SSD (1) with the cost function (4). Let the control input policy (38) and the parameter update law (37) be given. Let the initial value of the value function estimation i.e., $\hat{\theta}(0)$ be defined in a compact set such that the initial control input be admissible. Then, the estimated control input converges to the optimal one.

Proof. The control input error can be defined as $\tilde{u} = \hat{u} - u^*$ which gives $\tilde{u} = R^{-1} B^T \nabla^T \sigma(y)(\theta^* - \hat{\theta})$. Defining $\tilde{\theta} = \hat{\theta} - \theta^*$ and taking the norm of the control input error, one has $\|\tilde{u}\| = \|R^{-1} B^T \nabla^T \sigma(y)\| \|\tilde{\theta}\|$. It is shown in Theorem 2 that the parameter estimation error $\|\tilde{\theta}\|$ and the state of the system converge to zero asymptotically, if the PE condition is satisfied on the regression vector. Therefore, the estimated control input tends to the optimal control input as $t \to \infty$. ■

Next, the simulation results are presented.

5. SIMULATION RESULTS

A simulation example is provided, in this section, to demonstrate the efficacy of the proposed controller. Let an LTI-SSD be defined as

$$\begin{cases} \dot{x}_1(t) = 2x_1(t) + x_2(t - d_x) + u(t) \\ \dot{x}_2(t) = x_1(t) - x_2(t) \\ y_1 = 2x_1, \quad y_2 = x_1 + 1.2x_2 \end{cases}, \tag{61}$$

where

$$A_0 = \begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix}, A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

$$B = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, C = \begin{bmatrix} 2 & 0 \\ 1 & 1.2 \end{bmatrix}$$

The following profile is considered as the initial history of the state as

$$x_1(\theta) = x_2(\theta) = 1 \qquad \theta \in [-d_x, 0]\,. \tag{62}$$

The performance index is considered to be

$$J = \int_0^\infty \left(6y_1^2 + 2.4y_1y_2 + 2.44y_2^2 + u^2\right) dt\,, \tag{63}$$

which gives $\tilde{Q} = \begin{bmatrix} 6 & 1.2 \\ 1.2 & 2.44 \end{bmatrix}$, $R = 1$ with $\alpha = 1$ and $\gamma = 4$.

5.1. EVENT-TRIGGERED CONTROLLER WITH KNOWN INTERNAL DYNAMICS

In this subsection, the event-triggered controller is simulated assuming the internal dynamics of the system are known. Then, the solution to the OGDARE (16) becomes

$$P = \begin{bmatrix} 1.44 & -0.15 \\ -0.15 & 0.76 \end{bmatrix},$$

Using the design parameters, one can see that the condition provided in Lemma 1 holds as $\lambda_{min}(C^T C - I_2) = 0.1 > 0$. The output trajectories of the system (61) for different values of delays as $d_x = 0.001(s)$, $0.1(s)$, and $5(s)$ when the dynamics are known are shown in Figure 3. It can be seen from Figure 3 that as the delay in the system increases, it takes more time for the system to converge to the equilibrium point, which is expected. The convergence of the system to the equilibrium point and the output trajectory with respect to the cost function is shown in Figure 4.

5.2. EVENT-TRIGGERED ADAPTIVE CONTROL WITH UNKNOWN DRIFT DYNAMICS

In this subsection, it is assumed that the internal dynamics of the system (61), i.e., A_0, A, are considered unknown. To this end, the critic parameters such as regression vector is selected as follows

$$\sigma(y) = \begin{bmatrix} y_1^2 \\ y_1 y_2 \\ y_2^2 \end{bmatrix},$$

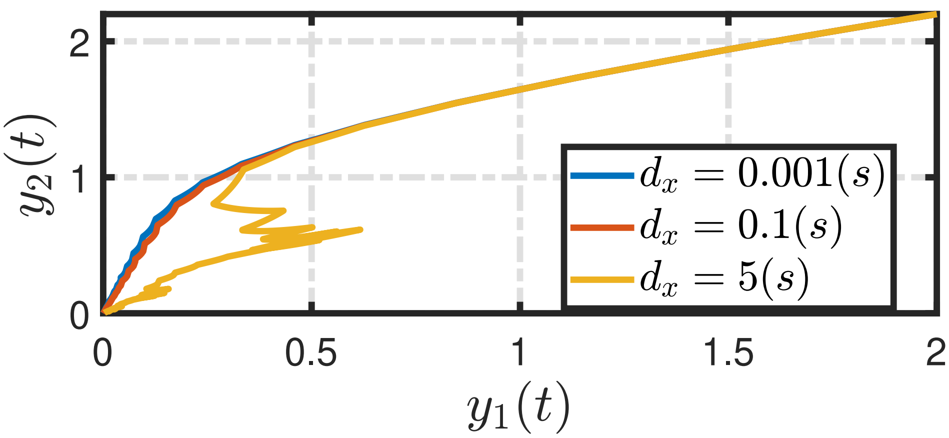


Figure 3. The output trajectories of the system (61) for different values of the state delay as $d_x = 0.001(s),\ 0.1(s),\ 5(s)$.

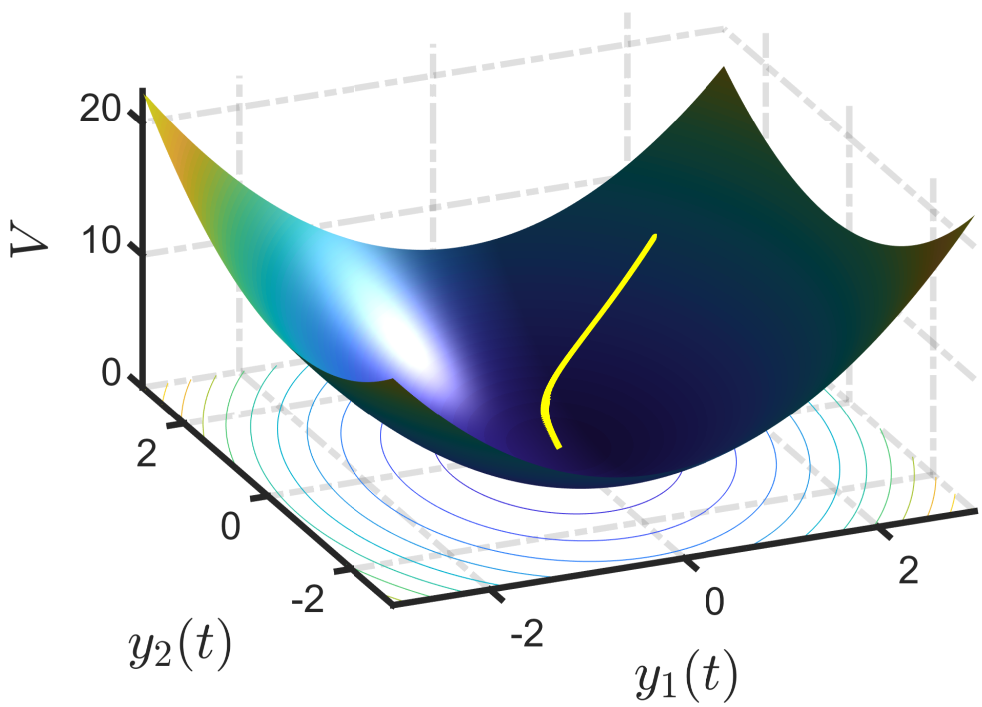


Figure 4. The optimal cost function surface when $d_x = 0.01(s)$.

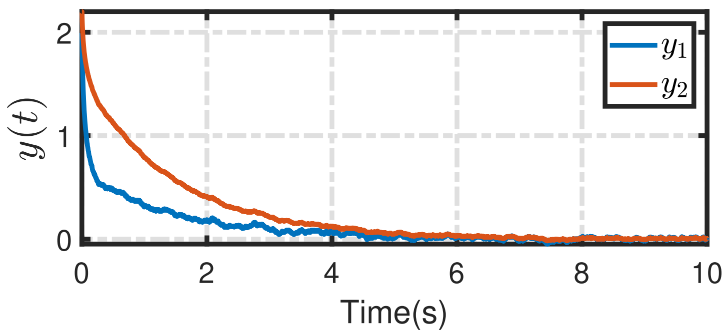
and the convergence rate $\beta = 0.85$. The performance of the hybrid RL scheme is shown in Figure 5 when the drift dynamics are unknown. An exploration random noise is added to the control input to excite all the modes of the system dynamics and to ensure the PE condition is fulfilled. This guarantees the convergence of the estimated parameters while the system states have not converged.

It is shown in Figure 5a that when the PE condition is removed the outputs of the system converge to zero. The effect of the PE condition on the output can be seen. The measurement error and event-trigger threshold is depicted in Figure 5b and the inter-event times is depicted in Figure 5c. It can be seen that there is not any Zeno behavior in the event-triggering mechanism, which shows the result of Theorem 2 for the case where the dynamics are not known.

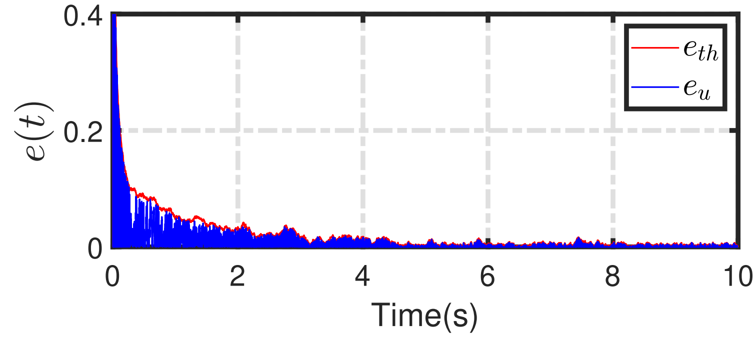
The convergence of the estimated parameters is shown in Figure 6. It can be seen that parameter estimation error is bounded and, when $t \to \infty$ and PE condition is hold, it is shown that the parameter estimation error converges to zero. This complies with the result of Theorem 3. Moreover, the integral temporal difference error (36) is demonstrated in Figure 7. One can see that the e_{ITDE} converges to zero before the PE condition holds which shows that the estimated value function converges to the optimal one. The output with respect to the cost function for both known and unknown dynamics is shown in Figure 8. It can be observed that for the unknown case initially the value of the cost function is high due to the initialization of the parameters, however, eventually as the parameters converges, the cost function also tends to zero.

6. CONCLUSION

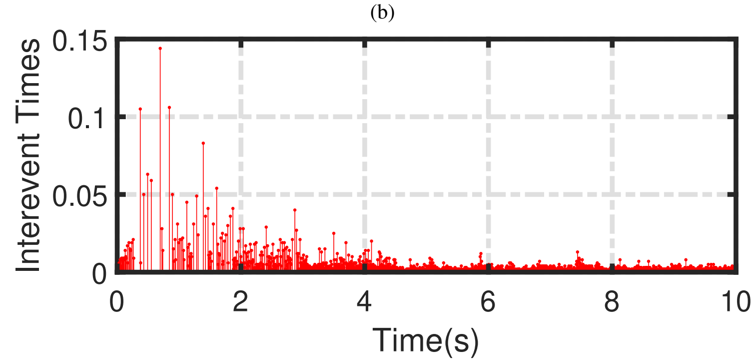
The event-triggered adaptive output feedback control of partially uncertain LTI-SSD by using IRL in the event-sampled formulation is addressed. The two player zero-sum game formulation of the time-delay system in the event-sampled framework appears to yield a saddle-point solution. A novel output delay game algebraic Riccati equation is needed and



(a)



(b)



(c)

Figure 5. Performance of the hybrid RL control scheme. (a) State vector, (b) Event threshold and measurement error, (c) Inter-event times

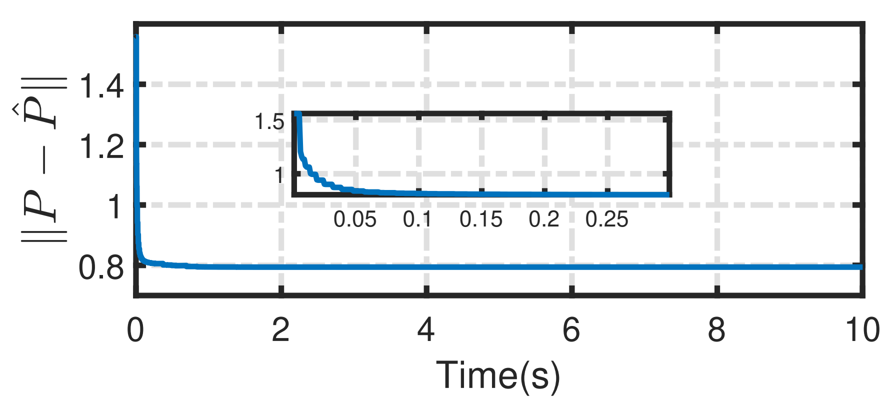


Figure 6. The parameter estimation error, i.e., $\|P - \hat{P}\|$.

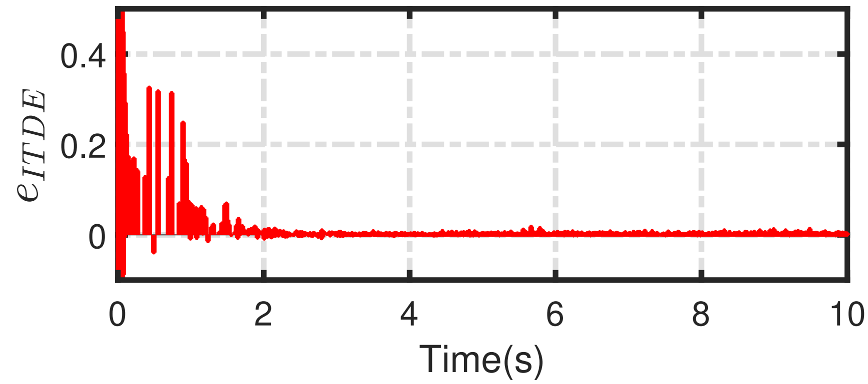


Figure 7. The integral temporal difference error, i.e., e_{ITDE}, (36).

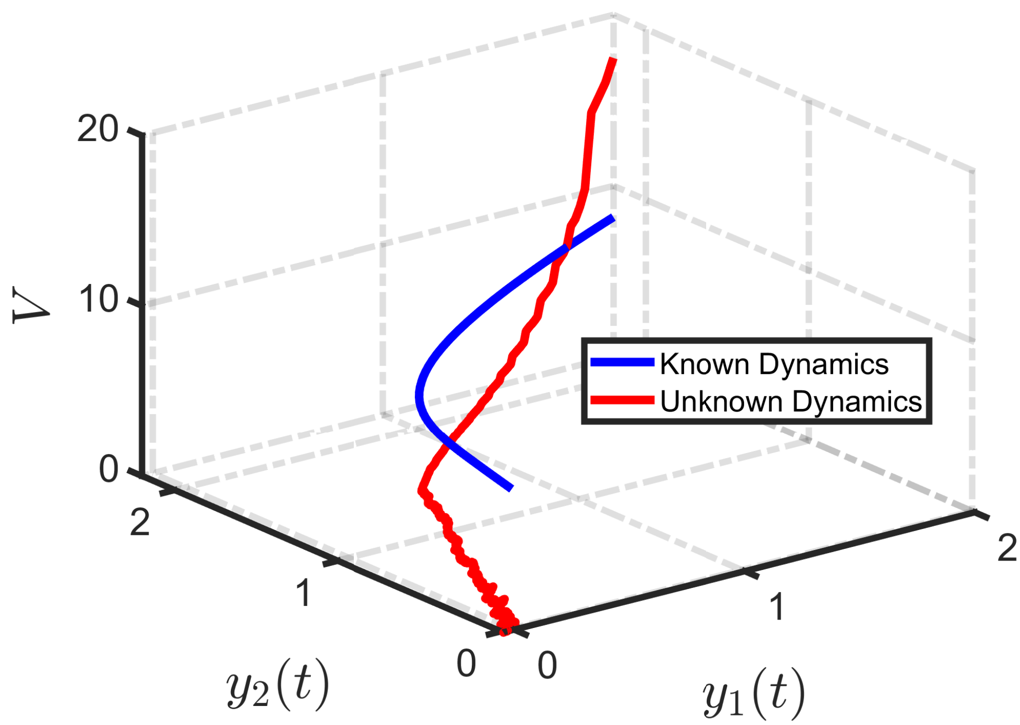


Figure 8. The output of the system vs cost function for both known and unknown dynamics.

derived given the system dynamics. It was demonstrated that when the system dynamics are known, the optimal control policy results in the asymptotic stability of the closed-loop system provided the linear time delayed system is controllable and observable. The event-triggered approach is used to relax the need of continuous availability of the system output.

Next, the IRL approach using actor-critic structure estimated the optimal control input and event-triggering threshold without requiring any information about drift dynamics and delayed state coefficient. However, to obtain the optimal control policy, the input matrix is still required. The hybrid learning scheme that is utilized to update the value function estimation in the event-sampled interval helps to attain the optimality faster. The asymptotic stability of the overall closed-loop system is shown using Lyapunov theory. It is shown that a PE condition is required to ensure the convergence of the value function parameters. A simulation example verified the effectiveness of the proposed methods. Using this approach, the authors conclude that event sampled framework helps to conceive the iterative and hybrid techniques needed for OAC.

REFERENCES

[1] J. P. Hespanha, P. Naghshtabrizi, and Y. Xu. A survey of recent results in networked control systems. Proceedings of the IEEE, 95(1):138–162, Jan 2007.

[2] Keqin Gu, Jie Chen, and Vladimir L Kharitonov. Stability of time-delay systems. Springer Science & Business Media, 2003.

[3] Emilia Fridman. Introduction to Time-delay Systems: Analysis and Control. Springer, 2014.

[4] Jean-Pierre Richard. Time-delay systems: an overview of some recent advances and open problems. Automatica, 39(10):1667 – 1694, 2003.

[5] Iasson Karafyllis, Michael Malisoff, Frederic Mazenc, and Pierdomenico Pepe. Recent results on nonlinear delay control systems. Springer International Publishing Switzerland, 2016.

[6] Y. Zhu, M. Krstic, and H. Su. Adaptive output feedback control for uncertain linear time-delay systems. IEEE Transactions on Automatic Control, 62(2):545–560, Feb 2017.

[7] A. Ponomarev. Reduction-based robustness analysis of linear predictor feedback for distributed input delays. IEEE Transactions on Automatic Control, 61(2):468–472, Feb 2016.

[8] Z. Artstein. Linear systems with delayed controls: A reduction. IEEE Transactions on Automatic Control, 27(4):869–879, August 1982.

[9] X. Cai, N. Bekiaris-Liberis, and M. Krstic. Input-to-state stability and inverse optimality of linear time-varying-delay predictor feedbacks. IEEE Transactions on Automatic Control, 63(1):233–240, Jan 2018.

[10] V. L. Kharitonov, S. . Niculescu, J. Moreno, and W. Michiels. Static output feedback stabilization: necessary conditions for multiple delay controllers. IEEE Transactions on Automatic Control, 50(1):82–86, Jan 2005.

[11] D. Bresch-Pietri and M. Krstic. Delay-adaptive predictor feedback for systems with unknown long actuator delay. IEEE Transactions on Automatic Control, 55(9):2106–2112, Sep. 2010.

[12] Miroslav Krstic. Delay compensation for nonlinear, adaptive, and PDE systems. Springer, 2009.

[13] Yang Zhu, Miroslav Krstic, and Hongye Su. Adaptive global stabilization of uncertain multi-input linear time-delay systems by pde full-state feedback. Automatica, 96:270 – 279, 2018.

[14] Kim-Doang Nguyen. An adaptive controller for uncertain nonlinear systems with multiple time delays. Automatica, 117:108976, 2020.

[15] Delphine Bresch-Pietri, Jonathan Chauvin, and Nicolas Petit. Adaptive control scheme for uncertain time-delay systems. Automatica, 48(8):1536 – 1552, 2012.

[16] H. Zhang, Y. Liu, G. Xiao, and H. Jiang. Data-based adaptive dynamic programming for a class of discrete-time systems with multiple delays. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 50(2):432–441, 2020.

[17] Y. Liu, H. Zhang, R. Yu, and Z. Xing. H_∞ tracking control of discrete-time system with delays via data-based adaptive dynamic programming. IEEE Transactions on Systems, Man, and Cybernetics: Systems, pages 1–8, 2019.

[18] D. Ross and I. Flügge-Lotz. An optimal control problem for systems with differential-difference equation dynamics. SIAM Journal on Control, 7(4):609–623, 1969.

[19] M. Basin and J. Rodriguez-Gonzalez. Optimal control for linear systems with multiple time delays in control input. IEEE Transactions on Automatic Control, 51(1):91–97, Jan 2006.

[20] D. Maity, M. H. Mamduhi, S. Hirche, K. H. Johansson, and J. S. Baras. Optimal lqg control under delay-dependent costly information. IEEE Control Systems Letters, 3(1):102–107, Jan 2019.

[21] F. Cacace, F. Conte, A. Germani, and G. Palombo. Optimal control of linear systems with large and variable input delays. Systems & Control Letters, 89:1 – 7, 2016.

[22] Richard S Sutton and Andrew G Barto. Reinforcement learning: An introduction. MIT press, 2018.

[23] F. L. Lewis and D. Vrabie. Reinforcement learning and adaptive dynamic programming for feedback control. IEEE Circuits and Systems Magazine, 9(3):32–50, 2009.

[24] V. Narayanan and S. Jagannathan. Event-triggered distributed control of nonlinear interconnected systems using online reinforcement learning with exploration. IEEE Transactions on Cybernetics, 48(9):2510–2519, Sep. 2018.

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

[26] D. P. Borgers, V. S. Dolk, and W. P. M. H. Heemels. Riccati-based design of event-triggered controllers for linear systems with delays. IEEE Transactions on Automatic Control, 63(1):174–188, Jan 2018.

[27] 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, 58(2):422–434, Feb 2013.

[28] V. Narayanan and S. Jagannathan. Event-triggered distributed approximate optimal state and output control of affine nonlinear interconnected systems. IEEE Transactions on Neural Networks and Learning Systems, 29(7):2846–2856, July 2018.

[29] I. Karafyllis and M. Krstic. Adaptive certainty-equivalence control with regulation-triggered finite-time least-squares identification. IEEE Transactions on Automatic Control, 63(10):3261–3275, Oct 2018.

[30] Dong Wang, Vijay Gupta, and Wei Wang. An event-triggered protocol for distributed optimal coordination of double-integrator multi-agent systems. Neurocomputing, 319:34 – 41, 2018.

[31] Vignesh Narayanan, Sarangapani Jagannathan, and Rohollah Moghadam. Optimality in event-triggered adaptive control of uncertain linear dynamical systems. In AIAA Scitech 2019 Forum, 2019.

[32] Kyriakos G. Vamvoudakis and Henrique Ferraz. Model-free event-triggered control algorithm for continuous-time linear systems with optimal performance. Automatica, 87:412 – 420, 2018.

[33] V. Narayanan, S. Jagannathan, and K. Ramkumar. Event-sampled output feedback control of robot manipulators using neural networks. IEEE Transactions on Neural Networks and Learning Systems, 30(6):1651–1658, 2019.

[34] O. Santos, S. Mondié, and V.L. Kharitonov. Linear quadratic suboptimal control for time delays systems. International Journal of Control, 82(1):147–154, 2009.

[35] A. Manitius and A. Olbrot. Finite spectrum assignment problem for systems with delays. IEEE Transactions on Automatic Control, 24(4):541–552, 1979.

[36] A. Sahoo, V. Narayanan, and S. Jagannathan. A min–max approach to event- and self-triggered sampling and regulation of linear systems. IEEE Transactions on Industrial Electronics, 66(7):5433–5440, 2019.

[37] Gilbert Strang. Introduction to Linear Algebra. Cambridge Press, 2016.

[38] Frank L Lewis, Draguna Vrabie, and Vassilis L Syrmos. Optimal Control. John Wiley & Sons, 2012.

[39] E. D. Sontag and Yuan Wang. New characterizations of input-to-state stability. IEEE Transactions on Automatic Control, 41(9):1283–1294, 1996.

[40] P. Pepe and Z.-P. Jiang. A lyapunov–krasovskii methodology for iss and iiss of time-delay systems. Systems & Control Letters, 55(12):1006 – 1014, 2006.

[41] Hassan K Khalil. Nonlinear control. Pearson New York, 2015.

IV. OPTIMAL ADAPTIVE CONTROL OF UNCERTAIN NONLINEAR CONTINUOUS-TIME SYSTEMS WITH INPUT AND STATE DELAYS

Rohollah Moghadam and S. Jagannathan

Department of Electrical and Computer Engineering

Missouri University of Science and Technology

Rolla, Missouri 65409–0050

Email: moghadamr@umsystem.edu

ABSTRACT

In this paper, a neural network (NN) based online optimal adaptive regulation of a class of nonlinear continuous-time systems with state and input delays with uncertain dynamics is introduced. The neural network (NN) based actor-critic framework estimates the cost function and optimal control input. The temporal difference error (TDE) is derived as a function of the difference between the actual and estimated cost function using integral reinforcement learning (IRL). The NN weights of the critic are tuned at every sampling instant as a function of the instantaneous integral TDE. A novel identifier is introduced to estimate the control coefficient matrices which are utilized in the generation of estimated control policy. The boundedness of the state vector, critic NN weights, identification error, and NN identifier weights are shown through Lyapunov analysis. Simulation results are provided to illustrate the effectiveness of the proposed approach.

1. INTRODUCTION

Time-delay systems have gained interest because of the importance of delay effects in practical applications, such as network-based control systems [1]. It is known that a delay in the feedback loop can affect the stability and performance of a physical system.

Therefore, the delay differential equations and the control of time-delay systems have been extensively studied in the literature [2–17]. Designing a controller that is optimal and robust against the system uncertainties for time-delay systems is of great importance.

The optimal regulation and tracking of time-delay systems has been reported in the literature [12, 18–22] when the system dynamics are known. An optimal control scheme is proposed in [19] for linear time-varying systems under multiple input delays by minimizing a quadratic cost function over finite time horizon. In [20], an optimal price-based decision making problem is addressed for a single-loop stochastic linear system in the presence of the network delay. These approaches [18–20, 23] are designed for linear systems and require system dynamics for the controller design, which is a bottleneck in many industrial systems.

To overcome the need for accurate knowledge of the system dynamics, adaptive control approaches [24] or reinforcement learning (RL) [25] can be utilized for obtaining optimal control policy for such time-delayed systems by solving the fixed-point Bellman or Hamilton-Jacobi-Bellman (HJB) equations. In particular, the IRL method [26] has been developed to generate optimal control of uncertain discrete and continuous-time systems in the absence of delays [27]. Given a cost function, the IRL utilizes the Bellman equation to obtain the optimal control policy online.

Optimal control of linear and nonlinear systems without delays using neural networks (NN) has been a commonly research area in the the past several decades [28, 29]. Theoretically, for an affine nonlinear system without delay, the optimal control policy over infinite horizon can be obtained by solving the Hamiltonian-Jacobi-Bellman (HJB) equation. However, the closed-form solution to the HJB equation is difficult to find [28, 30] and iterative techniques are typically employed. The value or policy iteration-based optimal adaptive techniques using NN-based adaptive dynamic programming (ADP) has been widely promoted in the control community for delay free systems [31–34]. The policy and/or value iteration-based schemes that utilize an initial stabilizing control input and cost/value function updates serve as a key technique to solve the HJB equation and to obtain optimal control policy. Due to the initial success, iterative optimal adaptive control using ADP has been studied for both continuous-time and discrete linear and nonlinear systems [32, 35, 36] using actor-critic framework which uses two NNs for approximating the value function and the control policy.

In contrast, for time-delay systems due to the infinite dimensional nature, optimal policy depends upon both delay and the history of the system state vector. It has been shown in [3] that state and input delays can cause instability to the overall closed-loop systems unless the effect of the delays is considered explicitly. In [18, 37], an additional integral term based on the solution to the differential equation is needed for computing the control policy, which is not preferred for real-time control. Also, the integral term causes pole-zero cancellation [38] when the system matrix is unstable. On the other hand, the finite dimensional, memoryless optimal control of time-delay systems with time-varying input delay is studied through traditional algebraic Ricatti equations (ARE) [23], provided a bound for the delay and control gain exists and complete knowledge of the system dynamics is known.

The design of the controller for nonlinear time-delay systems has also been addressed in the literature for both discrete-time and continuous-time systems [39–43]. In [39], an ADP approach is proposed to generate optimal control of nonlinear time-delay systems with state delay. On the other hand, the robust integral sliding mode for uncertain linear stochastic systems with time-varying delay is proposed in [40]. An adaptive NN control of nonlinear systems with unknown time delays is addressed in [41]. A continuous-time stabilizer is proposed for nonlinear systems with input and measurement delays in [43]. Although the control system design for nonlinear systems with time delays are investigated in the literature, the above mentioned schemes [39–43] consider either state delay [39, 41] or stabilizing control [40, 43]. To the best of our knowledge, the problem of finding the optimal adaptive NN control (OAC) of nonlinear time-invariant systems with state and input delays with partially or completely uncertain dynamics has not yet been reported.

In this paper, a finite dimensional and memoryless OAC of nonlinear continuous-time systems with input and state delays is introduced by using NN and ADP. It is assumed that the system dynamics are partially uncertain initially, that is, the control coefficient matrix is considered known. Then, the need for the control coefficient matrix is relaxed provided linear internal matrix of the non-delayed state and delay values are known. A quadratic cost function in terms of system state over an infinite time horizon is considered. Then, a Lyapunov-Krasovskii function consisting of system state and its delayed value is introduced as the value functional. The value function is then approximated using a NN and an integral term to incorporate the state delay.

Then, Hamiltonian function is formulated and, using the stationarity conditions, the optimal control is determined in terms of the system state and the critic NN weights when the dynamics are partially uncertain. Then, integral reinforcement learning (IRL) is used to estimate the value functional, named critic, through the integral temporal difference error (ITDE). The estimated NN weights update law is derived using the ITDE and the boundedness of the critic NN is discussed. Finally, using the estimated critic NN weight update law, the control policy, named Actor, is updated and the boundedness of the overall closed-loop system is provided.

Next, a novel NN identifier is introduced to estimate the system dynamics which is used to estimate the control coefficient matrix. Subsequently, the boundedness of the overall closed-loop system is analyzed with the NN identifier. Finally, simulation results are provided to show the effectiveness of the proposed approach.

The contributions of the paper include: 1) the selection of a value function in terms of input and state delays, 2) the definition of integral TDE (ITDE) incorporating delayed state vector, 3) the derivation of critic NN weight tuning law using ITDE, 4) the introduction of a NN identifier estimating the control coefficient matrix, 5) the development of finite dimensional memoryless OAC policy for nonlinear time-delayed systems including input and state delays under partial and completely uncertain dynamics, and 6) the Lyapunov stability analysis of the overall closed-loop system.

The paper is organized as follows. Background and problem statement are presented in Section 2. The optimal adaptive control of nonlinear time-delay systems with input and state delays is presented in Section 3. In Section 4, simulation results are given and, conclusions are provided in Section 5.

Notation. Throughout the paper, I_n is the $n \times n$ identity matrix and the Euclidean vector norm denotes $\|.\|$ denotes. The transpose of the matrix A is denoted as A^T. Maximum and minimum singular values of the matrix A are represented by $\delta_{max}(A)$ and $\delta_{min}(A)$, respectively. The minimum and maximum eigenvalue of the matrix A is defined as $\lambda_{min}(A)$ and $\lambda_{max}(A)$, respectively. The $ones(m, n)$ is a matrix with m rows and n columns in which all elements are equal to 1. The diagonal matrix of two matrices A and B is denoted as $diag\{A, B\}$.

2. BACKGROUND AND PROBLEM STATEMENT

In this section, the optimal control of a nonlinear continuous-time system is formulated. Consider the nonlinear continuous-time system described by

$$\begin{cases} \dot{x}(t) & = Ax(t) + f(x(t-d_x)) \\ & +g_1(x(t))u(x(t)) + g_2(x(t))u(x(t-d_u)) \\ x(\theta) & = \varphi_x(\theta) \qquad \theta \in [-d_x, 0] \\ u(\theta) & = \varphi_u(\theta) \qquad \theta \in [-d_u, 0] \end{cases}, \tag{1}$$

where $x(t) \in \Re^n$, $A \in \Re^{n\times n}$ is the linear internal dynamics of non-delayed state, $f(x(t-d_x)) \in \Re^n$ is a nonlinear function of delayed state vector, and the nonlinear input functions $g_1(x(t)).g_2(x(t)) \in \Re^{n\times m}$ satisfies $\|g_1(x(t))\|_F \leqslant g_{1M}$ and $\|g_2(x(t))\|_F \leqslant g_{2M}$ with $u(x(t)) \in \Re^m$ being the control input. The initial functions φ_x and φ_u are the history function and continuously differentiable.

Remark 1. The above continuous system has a linear matrix A which can be unstable. Typically, an n^{th} order affine nonlinear time-delay system in Brunovsky canonical form can be expressed as (1). For the optimal control policy development, the knowledge of the linear matrix A is not required. Moreover, note that the history function defined in (1) cannot be defined arbitrarily for the nonlinear systems because it has to satisfy the dynamics.

The objective is to generate the control policy in order to minimize the infinite horizon optimal value function defined as [28]

$$V(x(t)) = \int_0^\infty (x(\tau)^T Q x(\tau) + u(\tau)^T R u(\tau)) d\tau, \tag{2}$$

where $x(\tau)^T Q x(\tau) + u(\tau)^T R u(\tau)$ represents the local quadratic cost, $Q \in \Re^{n\times n}$ denotes a positive definite matrix and $R \in \Re^{m\times m}$ is a symmetric positive definite matrix.

Assumption 1. There exists a nominal linear model of the original nonlinear time-delay system which can be written as

$$\dot{x}(t) = A_0 x(t) + A_1 x(t-d_x) + B_0 u(t) + B_1 u(t-d_u)\,, \tag{3}$$

where A_0, A_1 are linear internal dynamic matrices and B_0 and B_1 are input coefficients matrices with appropriate dimensions.

Assumption 2. The nonlinear system (1) is controllable.

Remark 2. Assumptions 1 and 2 are not restrictive even for practical applications because for Assumption 1, apriori knowledge of the system dynamics is available such as an operating point. This nominal model around an operating point will aid in the design of the controller. Moreover, Assumption 2 is required to ensure that there exists a controller to stabilize the system.

Let the Hamiltonian function be defined as

$$H(x, u, V_x) = \frac{\partial V}{\partial t} + x(t)^T Q x(t) + u(t)^T R u(t) \tag{4}$$

where $\frac{\partial V}{\partial t}$ denotes the derivative of the value function with respect to time. To find the optimal policy for the nonlinear time-delay system (1) with the cost function (2), define a value functional as the output of a NN plus delayed state vector. Simply, let the optimal value functional be defined in terms of the NN weight matrix of the critic NN, system state and the delayed state as

$$\begin{aligned} V(x(t)) = W^T \sigma(x(t)) + \alpha_1 \int_{t-d_x}^{t} x^T x \, d\tau \\ + \alpha_2 \int_{t-d_u}^{t} x^T x \, d\tau + \epsilon(x), \end{aligned} \tag{5}$$

where W denotes the target NN weight matrix of the critic NN, $\sigma(x)$ is the activation function of the NN, α_1 and α_2 are scalar constants, and $\epsilon(x)$ is the NN function reconstruction error which is a function of state vector of the system.

Remark 3. To incorporate the effect of the delays in the control input, two integral terms are considered in the value function with respect to state and input delays. The proposed control policy can be applied to non-delay systems as well by assuming the design parameters $\alpha_1 = \alpha_2 = 0$.

Substituting the optimal value function defined in (5) into the Hamiltonian (4) to obtain the HJB equation

$$
\begin{aligned}
H &= W^T \nabla\sigma(x)\left(Ax(t) + f(x(t-d_x))\right. \\
&\quad \left. + g_1(x)u(t) + g_2(x)u(t-d_u)\right) + x^T Qx \\
&\quad + u^T Ru + (\alpha_1 + \alpha_2)x^T x \\
&\quad - \alpha_1 x^T(t-d_x)x^T(t-d_x) \\
&\quad - \alpha_2 x^T(t-d_u)x^T(t-d_u) + \epsilon'(x) = 0
\end{aligned} \quad . \tag{6}
$$

where $\epsilon'(x) = \partial\epsilon(x)/\partial t$. Next, using the necessary condition of optimality, the optimal control policy can be found by equating the derivative of the Hamiltonian function (4) with respect to the control input $u(t)$ to zero, i.e. $\partial H/\partial u(t) = 0$ which results in the following optimal control policy

$$
\begin{cases}
u^*(x(t)) = -R^{-1}G^T(x(t))(\nabla^T\sigma(x)W + \nabla^T\epsilon(x)) \\
G(x(t)) = g_1(x) + g_2(x)e^{A_0 d_u}
\end{cases} \tag{7}
$$

where $\nabla\sigma(x)$ and $\nabla\epsilon(x)$ are the gradients of the NN activation function and reconstruction error with respect to the system state. The gain matrix G is derived from two terms of the Hamiltonian function as $\partial g_1(x)u(t)/\partial u(t) + \partial g_2(x)u(t-d_u)/\partial u(t) = 0$. The second term has the derivative of the delayed input with respect to the current input. According to Assumption 1, it is assumed that a nominal model of the system is available and, therefore, the second derivative can be obtained from the same approach presented in [19] where $u(t-d_u)/\partial u(t) = e^{A_0 d_u}$ which by replacing the derivative term, gives the control coefficient matrix G. Note that as the number of neurons increases $\nabla\epsilon(x) \to 0$ [49].

Assumption 3. The NN function reconstruction error and its first time derivative are bounded [50], as a function of the state vector given by $\|\epsilon(x)\| \leqslant a_1\|x\|$ and $\|\epsilon'(x)\| \leqslant a_2\|x\|$ where a_1 and a_2 are known positive constants. Moreover, the target NN weights of the value function are bounded as $\|W\| \leqslant W_M$.

Remark 4. It can be noted from (6) that the HJB equation has additional terms with respect to the input and state delays further complicating its solution. Literature indicated that the closed-form solution to the traditional HJB equation with $\alpha_1 = \alpha_2 = 0$ is difficult to obtain and iterative techniques are proposed.

Remark 5. In this paper it is assumed initially that the input matrices g_1, g_2 and the nominal internal dynamic matrix A_0 are known and, therefore, the matrix G becomes known initially. However, the linear and nonlinear internal dynamics of the system, i.e. A and $f(x(t-d_x))$, are assumed to be unknown. Finding the optimal control policy, even when the control coefficient matrix is known, is still a challenge in the presence of state and input delays. Subsequently, the control coefficient matrix G is considered unknown and estimated through a NN identifier. The estimated matrix is used in the estimated optimal policy.

Next, the following fact is needed in order to proceed.

Fact. The closed-loop system is bounded above when the optimal control input is asserted such that, i.e., $Ax + \|f(x(t-d_x)) + g_1(x(t))u^*(x(t)) + g_2(x(t))u^*(x(t-d_u))\| \leqslant -\bar{k}\|x\|^2$ for a known constant $\bar{k}$. An upper bound for the optimal closed-loop system can be established using Lyapunov analysis.

The regulation of nonlinear continuous time-delay systems using NN and associated Lyapunov proof are discussed in the next section.

3. OPTIMAL ADAPTIVE REGULATION

In this section, an optimal adaptive NN-based controller is proposed for a nonlinear continuous time-delay system (1), first with known control coefficient matrix, and then considering it as unknown. To this end, a single layer NN with basis functions is utilized for approximating the value function. The tuning of the critic NN weights are provided using IRL while ensuring their boundedness. Subsequently, an NN identifier is introduced to relax the need for control coefficient matrix. The overall structure of the proposed approach is given in Figure 1.

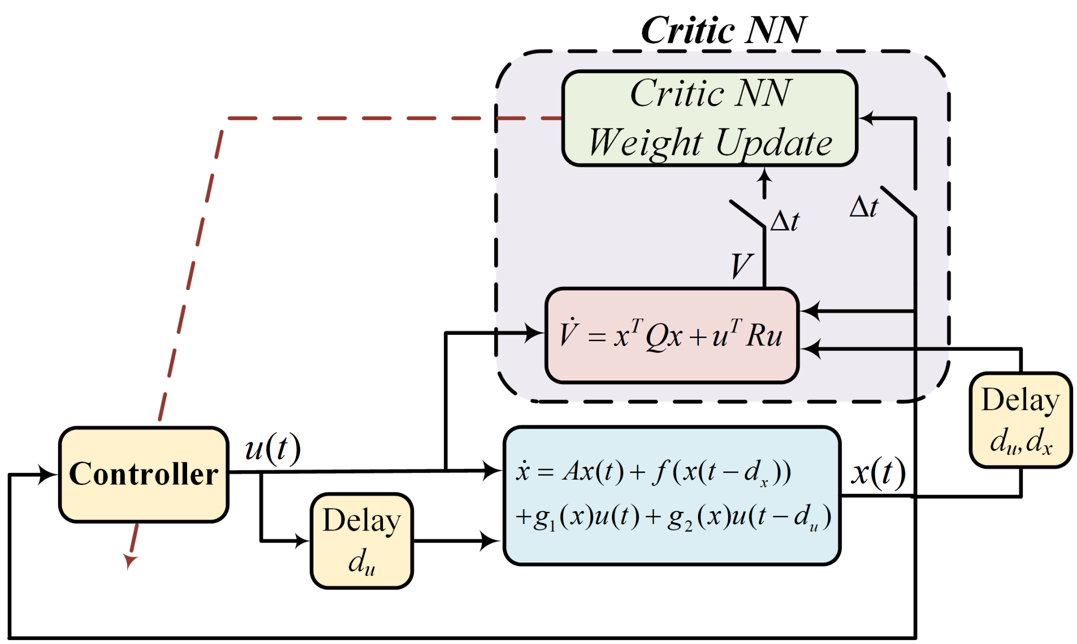


Figure 1. The overall structure of the proposed approach.

Notice that traditional dynamic programming works in a backward and off-line manner when the system dynamics are known, whereas an IRL approach helps in obtaining approximate optimal solutions in a forward-in-time manner even when the dynamics are uncertain. In contrast to using policy or value iterative techniques to solve the Bellman or HJB equation, the TDE is used as an alternate approach for finding the optimal control policy online and forward in time. The TDE can be treated as a prediction error between observed and predicted performance in response to the control policy when applied to the system. However, finding the solution to the Lyapunov equation arising from TDE makes it difficult to solve.

To overcome this issue, one can estimate value function by using a NN approximator the parameters of which can be adjusted by using ITDE. Then, a PE condition would be required to ensure the value function parameters converge asymptotically to the target optimal value.

In this paper, the ITDE is utilized to obtain OAC of a nonlinear time-delay system (1). The technique has two steps: critic and actor update. In IRL, the measured state vector at specific sampling times $t + i\Delta t$ with $i \in \mathbb{N}$ is used. The ITDE is employed to update the critic function with the current stabilizing controller. Then, the critic function is used to update the actor. The ITDE, which includes the history of the delayed state, is utilized to tune the value function weights under a PE condition.

By using the IRL, the Bellman equation can be written as

$$V(x(t)) = \int_{t}^{t+\Delta t} \left(x^T Q x + u^T R u\right) d\tau + V(x(t + \Delta t)) \,. \tag{8}$$

where $V(x)$ is the value function defined in (5). Now, the aim is to find the NN weights W in the value function which are utilized in the optimal control policy. Since due to the uncertainty in the dynamics of the system, the exact value of W cannot be determined, a value function approximator, called critic network, using a NN will be defined next and a tuning law to find the estimated weights of the value function, i.e. $\hat{W}$ will be introduced.

3.1. CRITIC NN APPROXIMATION

In this subsection, a NN is used to approximate the value function given in (5) and a tuning law is derived. Then, the boundedness of the critic network is shown using the Lyapunov analysis. To proceed, let the value function be approximated by

$$\begin{aligned}\hat{V}(x(t)) = \hat{W}^T\sigma(x(t)) + \alpha_1 \int_{t-d_x}^{t} x^T x d\tau \\ + \alpha_2 \int_{t-d_u}^{t} x^T x d\tau\end{aligned} \tag{9}$$

where $\hat{W}$ represents the estimated value of the actual value function weights, $\sigma(x(t)) : \Re^n \rightarrow \Re^N$ is the activation function vector with N being the number of hidden-layer neurons.

Remark 6. Since it is assumed that the state and also delayed state vector are available, the second and third terms in the value function (5) are known and, therefore, are not approximated.

Next, using the estimated value function, the following ITDE, i.e. E_{TD}, due to the difference between the actual and estimated parameters by using the Bellman equation can be defined as

$$\begin{aligned}E_{TD} =& \hat{V}(x(t+\Delta t)) - \hat{V}(x(t)) \\ &+ \int_{t}^{t+\Delta t} \left(x^T Q y + u^T R u\right) d\tau\end{aligned} \tag{10}$$

which can be written as

$$\begin{aligned}E_{TD} = &\int_{t}^{t+\Delta t} \left(x^T Q x + u^T R u\right) d\tau + \hat{W}^T \Delta\sigma(x) \\ &+ \alpha_1 \left(\int_{t+\Delta t-d_x}^{t+\Delta t} x^T(s)x(s)ds - \int_{t-d_x}^{t} x^T(s)x(s)ds\right) \\ &+ \alpha_2 \left(\int_{t+\Delta t-d_u}^{t+\Delta t} x^T(s)x(s)ds - \int_{t-d_u}^{t} x^T(s)x(s)ds\right)\end{aligned} \tag{11}$$

where $\Delta\sigma(x) = \sigma(x(t + \Delta t)) - \sigma(x(t))$.

Remark 7. The ITDE, defined in (11), is dependent upon state and input delays of the nonlinear system (1). In other words, the state and input delays influence the TDE and makes the proposed approach unique for the nonlinear time-delayed systems. This approach would be applicable to the delay free systems, if the coefficients of the integral terms are set to zero.

An appropriate update law should be designed for $\hat{W}$ such that the ITDE (11) tends to near zero, eventually. To this end, considering $E = \frac{1}{2}\|E_{TD}\|^2$ and using the gradient descend method, the following update law can be constructed as

$$\dot{\hat{W}} = -\beta \frac{\Delta\sigma(x(t))}{\left(1 + \Delta\sigma(x)^T \Delta\sigma(x)\right)^2} E_{TD}^T \tag{12}$$

where β is a scalar gain determining the convergence rate.

Remark 8. The update law for the NN weights defined in (12) is a function of the ITDE(11), which in turn is dependent upon delayed input and state vector. The critic NN weights are tuned by the current value of the state vector and by the delayed state and input vectors. Simply, the delays affect the learning of the value functional.

Next, define $\tilde{W} = W - \hat{W}$ as the weight estimation error and using (11) and (9), express the ITDE as a function of NN weight estimation as

$$E_{TD} = -\tilde{W}^T \Delta\sigma(x) - \Delta\epsilon(x) \tag{13}$$

where $\Delta\epsilon(x) = \epsilon(x(t + \Delta t)) - \epsilon(x(t))$, which is bounded above by $\|\Delta\epsilon(x)\| \leqslant \epsilon_{max}$. Then, the weight estimation error dynamics become

$$\dot{\tilde{W}} = \beta \frac{\Delta\sigma(x)}{\left(1 + \Delta\sigma(x)^T \Delta\sigma(x)\right)^2} E_{TD}^T \tag{14}$$

In the following theorem, the boundedness of the value functional is demonstrated by using an initial stabilizing control input $u(x)$. The nonlinear system dynamics must remain stable as a result of applying an admissible control policy.

Theorem 1. (Boundedness of the critic NN) Consider the nonlinear continuous-time system (1) along with the value functional (2). Let Assumptions 1-3 be satisfied and the initial control policy $u(t)$ be admissible and the critic NN weights update law be given by (12) and $\frac{\Delta\sigma(x)}{(1+\Delta\sigma(x)^T\Delta\sigma(x))^2}$ be persistently exciting. Then, the critic weight estimation error is ultimately bounded.

Proof. Consider the positive Lyapunov candidate function

$$V_1 = \frac{1}{2}tr\left\{\tilde{W}^T\tilde{W}\right\} \tag{15}$$

Using the weight estimation error update law (14), the first derivative of the Lyapunov function candidate (15), i.e. $\dot{V}_1 = tr\left\{\tilde{W}^T\dot{\tilde{W}}\right\}$, is written as

$$\begin{aligned}\dot{V}_1 = &-\beta tr\left\{\tilde{W}^T\frac{\Delta\sigma(x)\Delta\sigma(x)^T}{(1+\Delta\sigma(x)^T\Delta\sigma(x))^2}\tilde{W}\right\}\\ &+\beta tr\left\{\tilde{W}^T\frac{\Delta\sigma(x)\Delta\epsilon(x)^T}{(1+\Delta\sigma(x)^T\Delta\sigma(x))^2}\right\}\end{aligned} \tag{16}$$

Using the trace properties of matrices and after simplifications, (16) becomes

$$\begin{aligned}\dot{V}_1 \leqslant &-\beta\left\|\frac{\Delta\sigma(x)^T}{(1+\Delta\sigma(x)^T\Delta\sigma(x))}\tilde{W}\right\|^2\\ &+\beta\left\|\frac{\Delta\sigma(x)^T}{(1+\Delta\sigma(x)^T\Delta\sigma(x))}\tilde{W}\right\|\left\|\frac{\Delta\epsilon(x)^T}{(1+\Delta\sigma(x)^T\Delta\sigma(x))}\right\|\end{aligned} \tag{17}$$

which can be expressed by

$$\dot{V}_1 \leqslant -\beta \left\| \frac{\Delta\sigma(x)^T}{(1+\Delta\sigma(x)^T\Delta\sigma(x))}\tilde{W} \right\| \left(\left\| \frac{\Delta\sigma(x)^T}{(1+\Delta\sigma(x)^T\Delta\sigma(x))}\tilde{W} \right\| - \left\| \frac{\Delta\epsilon(x)^T}{(1+\Delta\sigma(x)^T\Delta\sigma(x))} \right\| \right) \tag{18}$$

The first derivative of the Lyapunov candidate function (15), given in (18) becomes less than zero, i.e. $\dot{V}_1 < 0$, if and only if $\left\| \frac{\Delta\sigma(x)^T}{(1+\Delta\sigma(x)^T\Delta\sigma(x))}\tilde{W} \right\| > \left\| \frac{\Delta\epsilon(x)^T}{(1+\Delta\sigma(x)^T\Delta\sigma(x))} \right\|$. Since the NN approximation error is considered bounded, then, one can conclude $\left\| \frac{\Delta\epsilon(x)^T}{(1+\Delta\sigma(x)^T\Delta\sigma(x))} \right\| \leqslant \bar{\epsilon}$. Consequently, using the bounds one can infer that the weight estimation error of the critic NN is bounded and, therefore, the critic NN is bounded provided the PE condition holds. ■

Remark 9. Note that when the number of the neurons in the hidden layer increases, i.e. $N \to \infty$, the approximation error converges to zero [49] and, therefore, the bound for $\Delta\epsilon(x)$ decreases.

Next, the approximated NN weights of the Critic NN are utilized in the optimal control policy under the assumption that the control coefficient matrix is known. The boundedness of the overall closed-loop system is shown using Lyapunov theory.

3.2. CONTROL POLICY AND OVERALL STABILITY WITH KNOWN G MATRIX

In this subsection, the control coefficient matrix is considered known. By using the known control coefficient matrix, the optimal policy is generated as

$$\hat{u}(x(t)) = -R^{-1}G^T\nabla^T\sigma(x)\hat{W} \tag{19}$$

In the following theorem, the estimated critic NN weights along with the estimated control input are incorporated into the overall closed-loop system and the boundedness of the overall closed-loop system is given using the Lyapunov theory.

Theorem 2. Consider the nonlinear continuous time-delay system (1) along with the cost function (2). Let Assumptions 1-3 be satisfied and the initial control policy $u(t)$ be admissible. Consider the critic NN weights update law as (12) and let $\frac{\Delta\sigma(x)}{(1+\Delta\sigma(x)^T\Delta\sigma(x))^2}$ be persistently exciting. Let the design parameters be designed such that $\alpha_1 > 0$ and $\alpha_2 > 0$ and $0 < \alpha_1 + \alpha_2 < (\bar{k} + \bar{\epsilon})$, where $\bar{k}$ and $\bar{\epsilon}$ represent a bound for the system under actual optimal control input and NN approximation error bound respectively. Then, the overall closed-loop system remains ultimately bounded and the estimated optimal control policy becomes bounded close to the optimal policy.

Proof. Consider the positive Lyapunov candidate function as

$$L = V(x) + \frac{1}{2}tr\left\{\tilde{W}^T\tilde{W}\right\} \tag{20}$$

where $V(x)$ is given in (5). Taking the first derivative of (20) to get

$$\dot{L} = \dot{V}(x) + tr\left\{\tilde{W}^T\dot{\tilde{W}}\right\} \tag{21}$$

The derivative of the first term of the overall Lyapunov candidate function can be written as

$$\begin{aligned}\dot{V} &= W^T(\nabla\sigma(x)\dot{x}) + (\alpha_1+\alpha_2)x^Tx \\ &- \alpha_1 x^T(t-d_x)x(t-d_x) \\ &- \alpha_2 x^T(t-d_u)x(t-d_u) + \nabla\epsilon(x)\dot{x}\end{aligned} \tag{22}$$

which by substituting the system dynamics (1) can be expressed as

$$
\begin{aligned}
\dot{V} &= W^T \nabla\sigma(x)(Ax(t) + f(x(t-d_x)) \\
&+ g_1\hat{u} + g_2\hat{u}(x(t-d_u))) + (\alpha_1 + \alpha_2)x^T x \\
&- \alpha_1 x^T(t-d_x)x(t-d_x) - \alpha_2 x^T(t-d_u)x(t-d_u) \\
&+ \nabla\epsilon(x)(Ax(t) + f(x(t-d_x)) \\
&+ g_1\hat{u} + g_2\hat{u}(x(t-d_u)))
\end{aligned} \tag{23}
$$

Substituting the estimated control input (19) into (23) yields

$$
\begin{aligned}
\dot{V} &= W^T \nabla\sigma(Ax(t) + f(x(t-d_x))) \\
&- g_1 W^T \nabla\sigma R^{-1} G^T \nabla^T \sigma \hat{W} \\
&- g_2 W^T \nabla\sigma R^{-1} G^T \nabla^T \sigma(x(t-d_u))\hat{W} \\
&- \alpha_1 x^T(t-d_x)x(t-d_x) - \alpha_2 x^T(t-d_u)x(t-d_u) \\
&+ (\alpha_1 + \alpha_2)x^T x + \nabla\epsilon(Ax(t) + f(x(t-d_x))) \\
&- g_1 W^T \nabla\epsilon R^{-1} G^T \nabla^T \sigma \hat{W} \\
&- g_2 W^T \nabla\epsilon R^{-1} G^T \nabla^T \sigma(x(t-d_u))\hat{W}
\end{aligned} \tag{24}
$$

Now, replacing the weight estimation error defined as $\hat{W} = W - \tilde{W}$ in (24), one gets

$$
\begin{aligned}
\dot{V} &= W^T \nabla\sigma(Ax(t) + f(x(t-d_x))) \\
&- g_1 W^T \nabla\sigma R^{-1} G^T \nabla^T \sigma(W - \tilde{W}) \\
&- g_2 W^T \nabla\sigma R^{-1} G^T \nabla^T \sigma(x(t-d_u))(W - \tilde{W}) \\
&- \alpha_1 x^T(t-d_x)x(t-d_x) - \alpha_2 x^T(t-d_u)x(t-d_u) \\
&+ \nabla\epsilon(Ax(t) + f(x(t-d_x))) + (\alpha_1 + \alpha_2)x^T x \\
&- g_1 W^T \nabla\epsilon R^{-1} G^T \nabla^T \sigma(W - \tilde{W}) \\
&- g_2 W^T \nabla\epsilon R^{-1} G^T \nabla^T \sigma(x(t-d_u))(W - \tilde{W})
\end{aligned} \tag{25}
$$

which, after some simplifications by separating the terms with respect to the actual NN weights and the terms with respect to the weight estimation error, (25) can be rewritten as

$$\begin{aligned}
\dot{V} &= W^T \nabla\sigma(Ax(t) + f(x(t-d_x)) \\
&- g_1 W^T \nabla\sigma R^{-1} G^T \nabla^T \sigma W \\
&- g_2 W^T \nabla\sigma R^{-1} G^T \nabla^T \sigma(x(t-d_u))W \\
&- \alpha_1 x^T(t-d_x)x(t-d_x) - \alpha_2 x^T(t-d_u)x(t-d_u) \\
&+ \nabla\epsilon(Ax(t) + f(x(t-d_x)) + (\alpha_1+\alpha_2)x^T x \\
&- g_1 W^T \nabla\epsilon R^{-1} G^T \nabla^T \sigma W \\
&- g_2 W^T \nabla\epsilon R^{-1} G^T \nabla^T \sigma(x(t-d_u))W \\
&+ g_1 W^T \nabla\sigma R^{-1} G^T \nabla^T \sigma \tilde{W} \\
&+ g_2 W^T \nabla\sigma R^{-1} G^T \nabla^T \sigma(x(t-d_u))\tilde{W} \\
&+ g_1 W^T \nabla\epsilon R^{-1} G^T \nabla^T \sigma \tilde{W} \\
&+ g_2 W^T \nabla\epsilon R^{-1} G^T \nabla^T \sigma(x(t-d_u))\tilde{W}
\end{aligned} \tag{26}$$

Now, using Fact 1 and assuming the nonlinear system with actual optimal control input is bounded, i.e. $W^T\nabla\sigma(Ax(t) + f(x(t-d_x)) - g_1 W^T\nabla\sigma R^{-1}G^T\nabla^T\sigma W - g_2 W^T\nabla\sigma R^{-1}G^T\nabla^T\sigma(x(t-d_u))W \leqslant -\bar{k}\|x\|^2$. Moreover, the effect of the NN approximation error in the system dynamics is also assumed to be bounded above, i.e. $\nabla\epsilon(Ax(t) + f(x(t-d_x)) - g_1 W^T\nabla\epsilon R^{-1}G^T\nabla^T\sigma W - g_2 W^T\nabla\epsilon R^{-1}G^T\nabla^T\sigma(x(t-d_u))W \leqslant \bar{\epsilon}\|x\|^2$. Then, using the defined bounds and after a few manipulations, (26) can be expressed as

$$\begin{aligned}
\dot{V} &\leqslant -(\bar{k} + \bar{\epsilon} - \alpha_1 - \alpha_2)\|x\|^2 \\
&- \alpha_1\|x(t - d_x)\|^2 - \alpha_2\|x(t - d_u)\|^2 \\
&+ \big(g_1 W^T \nabla\sigma R^{-1} G^T \nabla^T \sigma \\
&+ g_2 W^T \nabla\sigma R^{-1} G^T \nabla^T \sigma(x(t - d_u))\big)\, \tilde{W} \\
&+ \big(g_1 W^T \nabla\epsilon R^{-1} G^T \nabla^T \sigma \\
&+ g_2 W^T \nabla\epsilon R^{-1} G^T \nabla^T \sigma(x(t - d_u))\big)\, \tilde{W}
\end{aligned} \tag{27}$$

Next, assuming the coefficients1 of the weight estimation error are bounded due to the bounded actual NN weights as $g_1 W^T \nabla\sigma R^{-1} G^T \nabla^T \sigma + g_2 W^T \nabla\sigma R^{-1} G^T \nabla^T \sigma(x(t - d_u)) \leqslant \bar{k}_1$ and $g_1 W^T \nabla\epsilon R^{-1} G^T \nabla^T \sigma + g_2 W^T \nabla\epsilon R^{-1} G^T \nabla^T \sigma(x(t - d_u)) \leqslant \bar{k}_2$, (27) turns into

$$\begin{aligned}
\dot{V} &\leqslant -(\bar{k} + \bar{\epsilon} - \alpha_1 - \alpha_2)\|x\|^2 \\
&- \alpha_1\|x(t - d_x)\|^2 - \alpha_2\|x(t - d_u)\|^2 \\
&+ (\bar{k}_1 + \bar{k}_2)\|\tilde{W}\|
\end{aligned} \tag{28}$$

The derivative of the second term of the overall Lyapunov candidate function is given in the proof of Theorem 1 as (17). Collecting the derivatives of the two corresponding terms in the overall Lyapunov candidate function under the PE condition, the first derivative of the Lyapunov candidate function can be written as

$$\begin{aligned}
\dot{L} &\leqslant -(\bar{k} + \bar{\epsilon} - \alpha_1 - \alpha_2)\|x\|^2 \\
&- \alpha_1\|x(t - d_x)\|^2 - \alpha_2\|x(t - d_u)\|^2 \\
&+ (\bar{k}_1 + \bar{k}_2)\|\tilde{W}\| - \beta \left\|\tilde{W}\right\|^2 + \beta\epsilon_{max}\|\tilde{W}\|
\end{aligned} \tag{29}$$

Defining $\rho = \epsilon_{max} + \frac{\bar{k}_1+\bar{k}_2}{\beta}$ and after some simplifications, (29) becomes

$$
\begin{aligned}
\dot{L} &\leqslant -(\bar{k} + \bar{\epsilon} - \alpha_1 - \alpha_2)\|x\|^2 \\
&- \alpha_1\|x(t - d_x)\|^2 - \alpha_2\|x(t - d_u)\|^2 \\
&- \beta\left(\left(\|\tilde{W}\| - \frac{\rho}{2}\right)^2 - \frac{\beta\rho^2}{4}\right)
\end{aligned}
\tag{30}
$$

It can be observed from (30) that the overall Lyapunov candidate function L becomes less than zero, i.e. $\dot{L} < 0$ if $\alpha_1 > 0$ and $\alpha_2 > 0$ and $0 < \alpha_1 + alpha_2 < (\bar{k} + \bar{\epsilon})$ provided one of the following conditions hold

$$
\begin{cases}
\|x\| > \dfrac{\sqrt{\beta}\rho}{2\sqrt{(\bar{k} + \bar{\epsilon} - \alpha)}} \\
\|x(t - d_x)\| > \dfrac{\sqrt{\beta}\rho}{2\sqrt{\alpha_1}} \\
\|x(t - d_u)\| > \dfrac{\sqrt{\beta}\rho}{2\sqrt{\alpha_2}} \\
\|\tilde{W}\| > \rho
\end{cases}
$$

To show the boundedness of the optimal control policy, the control input error can be defined as $\tilde{u} = u - \hat{u}$. Substituting the actual control input (7) and the estimated control input (19), one has $\tilde{u} = -R^{-1}G^T\nabla^T\sigma(x)(W - \hat{W}) - R^{-1}G^T(x(t))\nabla\epsilon(x)$. By using the weight estimation error $\tilde{W} = W - \hat{W}$, the control input error becomes $\tilde{u} = -R^{-1}G^T\nabla^T\sigma(x)\tilde{W} - R^{-1}G^T(x(t))\nabla\epsilon(x)$. Since the state vector is bounded, then, $\nabla\epsilon(x)$ becomes bounded. Moreover, as mentioned in [49] when the number of neurons increases $\nabla\epsilon(x) \rightarrow 0$. Therefore, boundedness of the state vector and NN weight estimation errors results in the boundedness of the control input error and as a consequence, the estimated optimal control input is bounded close to the optimal control policy. ■

Remark 10. For the boundedness of the overall closed-loop system, the bounds are required for the delayed states corresponding to input and state delays as they appear in the derivative of the Lyapunov function. To ensure the Lyapunov function is less than zero, the history of the states have to be selected in a specific bound. Moreover, bounds are functions of the design parameters given in the definition of the value function.

3.3. CONTROL POLICY WITH UNKNOWN G MATRIX

In this subsection, a NN to approximate the system internal dynamics and a second NN identifier to generate the estimated control coefficient matrix are utilized. Then, the estimated control policy uses the estimated $\hat{G}$ which relaxes the need to know the control coefficient matrix. Addition of an identifier using a NN complicates the overall closed-loop stability analysis. However, it is shown using the Lyapunov analysis, that the overall closed-loop system remains bounded in the presence of critic NN and the estimated control coefficient matrix. First, the online identifier to generate the estimated control coefficient matrix,$\hat{G}$, will be presented. Subsequently, the estimated control policy using the identifier is introduced and then, the overall stability analysis is discussed.

3.3.1. ONLINE SYSTEM IDENTIFIER. An online neural network based identifier is constructed for the nonlinear time-delay system (1) with input and state delay which will be used later in computing the control coefficient matrix G in (7).

To proceed, the nonlinear time-delay system (1) can be rewritten as

$$\dot{x}(t) = \begin{bmatrix} Ax(t) & 0 & 0 & 0 \\ 0 & f(x(t-d_x)) & 0 & 0 \\ 0 & 0 & g_1(x(t)) & 0 \\ 0 & 0 & 0 & g_2(x(t)) \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ u(t) \\ u(t-d_u) \end{bmatrix} \tag{31}$$

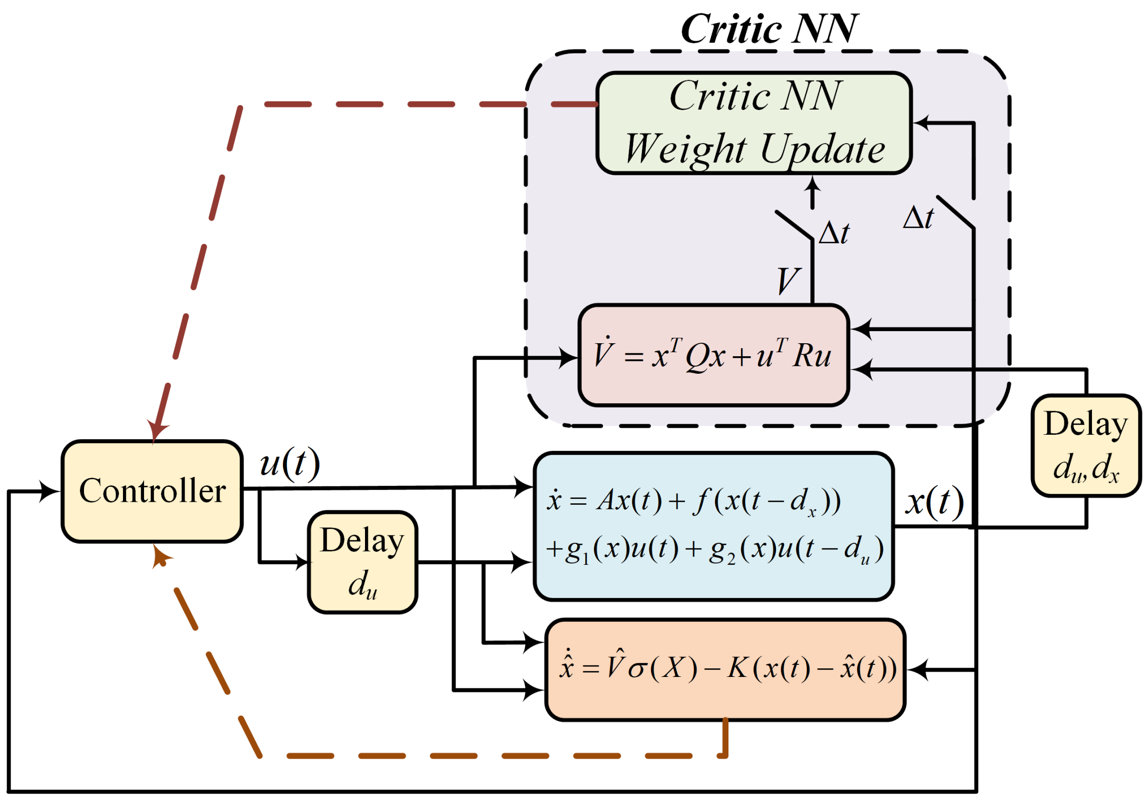


Figure 2. The overall structure of the proposed approach when the gain matrix G is unknown.

Defining the augmented control input as $\bar{u} = \begin{bmatrix} 1 & 1 & u(t) & u(t-d_u) \end{bmatrix}^T \in \Re^{2m+2}$ and using the universal function approximation property of the NN, the nonlinear system (31) can be represented on a compact set as $Ax = V_A^T \sigma_A(x) + \varepsilon_A(x)$, $f(x(t-d_x)) = V_f^T \sigma_f(x(t-d_x)) + \varepsilon_f(x)$, $g_1(x) = V_{g_1}^T \sigma_{g_1}(x) + \varepsilon_{g_1}(x)$, and $g_2(x) = V_{g_2}^T \sigma_{g_2}(x) + \varepsilon_{g_2}(x)$ where $V_A \in \Re^{l\times n}$, $V_f \in \Re^{l\times n}$, $V_{g_1} \in \Re^{l\times n}$, $V_{g_2} \in \Re^{l\times n}$ denote the target NN weight matrices and $\sigma_A(x) \in \Re^l$, $\sigma_f(x(t-d_x)) \in \Re^l$, $\sigma_{g_1}(x) \in \Re^{l\times m}$, $\sigma_{g_2}(x) \in \Re^{l\times m}$ represent the activation functions and $\varepsilon_A \in \Re^n$, $\varepsilon_f \in \Re^n$, $\varepsilon_{g_1} \in \Re^{n\times m}$, $\varepsilon_{g_2} \in \Re^{n\times m}$ are the NN reconstruction errors, respectively. Now, replacing the NN approximators in the nonlinear system (31) gives

$$\dot{x}(t) = \begin{bmatrix} V_A \\ V_f \\ V_{g_1} \\ V_{g_2} \end{bmatrix}^T \begin{bmatrix} \sigma_A(x) & 0 & 0 & 0 \\ 0 & \sigma_f(x(t-d_x)) & 0 & 0 \\ 0 & 0 & \sigma_{g_1}(x) & 0 \\ 0 & 0 & 0 & \sigma_{g_2}(x) \end{bmatrix} \bar{u} + \varepsilon_I(x) \tag{32}$$

which can be written as

$$\dot{x}(t) = V^T \boldsymbol{\sigma}(\boldsymbol{X})\bar{u} + \varepsilon_I(x) \tag{33}$$

where $V = \begin{bmatrix} V_A^T & V_f^T & V_{g_1}^T & V_{g_2}^T \end{bmatrix}^T \in \Re^{4l\times n}$ represents the augmented NN identifier weights and $\boldsymbol{\sigma}(\boldsymbol{X}) = diag\{\sigma_A(x), \sigma_f(x(t-d_x)), \sigma_{g_1}(x), \sigma_{g_2}(x)\} \in \Re^{4l\times(2m+2)}$ denotes the augmented NN identifier activation function. The NN identifier reconstruction error $\varepsilon_I(x)$ is defined as $\varepsilon_I(x) = \varepsilon_A(x) + \varepsilon_f(x) + \varepsilon_{g_1}(x)u + \varepsilon_{g_2}(x)u(t-d_u)$. Note that, since $\varepsilon_I(x)$ is a function of the system state $x(t)$, input $u(t)$ and delayed input $u(t-d_u)$, in this paper it is not considered that $\varepsilon_I(x)$ is bounded above by a constant value. Instead, it is assumed that it is bounded with respect to the norm of the state and its delayed values.

Assumption 4. The NN identifier reconstruction error is considered to be bounded as $\|\varepsilon_I(x)\|^2 \leqslant b_0\|x\|^2 + b_1\|x(t-d_x)\|^2 + b_3\|x(t-d_u)\|^2$ where b_0, b_1 and b_2 are known positive constants [50]. Besides, the target NN weights of the identifier are bounded as $\|V\| \leqslant V_M$.

It should be noted that since the NN identifier augmented activation function $\boldsymbol{\sigma}(\boldsymbol{X})$ is known, a proper tuning law for the NN identifier weight matrix V results in an appropriate approximation of the nonlinear time-delay system dynamics which will be used in finding the control input coefficient matrix. To this end, let the dynamics of the state estimator\identifier be defined as

$$\dot{\hat{x}}(t) = \hat{V}^T \boldsymbol{\sigma}(\boldsymbol{X})\bar{u} + K(x - \hat{x}) \tag{34}$$

where $\hat{x}(t) \in \Re^n$ is the state of the estimator and $K \in \Re^{n\times n}$ is a design gain matrix. Define the state estimation error as

$$e = x - \hat{x} \tag{35}$$

Next, using the dynamics of the nonlinear time-delay system (33) and the state estimator (34), the dynamics of the state estimation error can be written as

$$\dot{e} = -Ke + \tilde{V}\boldsymbol{\sigma}(\boldsymbol{X})\bar{u} + \varepsilon_I(x) \tag{36}$$

where $\tilde{V} = V - \hat{V}$ is the NN identifier weight error. Note that the state estimation error dynamics of the nonlinear time-delay system is a function of augmented input vector. Using the state estimation error (35), to enforce the convergence of the weight estimation error $\tilde{V}$ and the state estimation error e, the following update law is proposed for the NN identifier weight as

$$\dot{\hat{V}} = -\alpha_v \hat{V} + \boldsymbol{\sigma}(\boldsymbol{X})\bar{u}\, e^T \tag{37}$$

where $\alpha_v > 0$ is the tuning parameter, $\bar{u}$ and e are augmented control input and state estimation error, respectively. This implies that the identifier NN weights are tuned as a product of augmented input vector and state estimation error.

Then, the dynamics of the NN identifier weight estimation error $\tilde{V}$ becomes

$$\dot{\tilde{V}} = -\alpha_v \tilde{V} - \boldsymbol{\sigma}(\boldsymbol{X})\bar{u}e^T + \alpha_v V \tag{38}$$

The estimated NN identifier weights will be employed in the next section to find the estimated coefficient control input matrix G and results in the estimated control policy.

3.3.2. ESTIMATED CONTROL POLICY. The estimated input coefficient matrices defined in the previous subsection is used now to construct the control coefficient gain matrix G and relax the need for the input matrix and linear model given in Assumption 1. To this end, let the estimated optimal control input be given by

$$\hat{u}(t) = -R^{-1}\hat{G}^T \nabla^T \sigma(x)\hat{W} \tag{39}$$

where $\hat{G} = \hat{g}_1 + \hat{g}_2 e^{A_0 d_u}$. The estimated control input is now applied to the system and the boundedness of the overall closed-loop system is shown in the following theorem.

Theorem 3. Consider the nonlinear continuous time-delay system (1) along with the cost function (2) under the estimated control input (39). Let Assumptions 1-4 be satisfied and the state identifier (34) be defined with the weight update law (37). Let the initial control policy $u(t)$ be admissible and the critic NN weights update law be given by (12) and $\frac{\Delta\sigma(x)}{(1+\Delta\sigma(x)^T\Delta\sigma(x))^2}$ be persistently exciting where $\bar{k}$ and $\bar{\epsilon}$ represent a bound for the system under actual optimal control input and NN approximation error bound, respectively. Then, the overall closed-loop system remains ultimately bounded. Next, the estimated control policy is bounded close to the optimal policy.

Proof. Consider the positive Lyapunov candidate function as

$$\begin{aligned} L &= V_1 + V_2 + V_3 + V_4 \\ &= V(x) + \frac{1}{2}tr\left\{\tilde{W}^T\tilde{W}\right\} + \frac{1}{2}e^T e + \frac{1}{2}tr\left\{\tilde{V}^T\tilde{V}\right\} \end{aligned} \tag{40}$$

where $V(x)$ is given in (5). Taking the first derivative of (40), one has

$$\dot{L} = \dot{V}(x) + tr\left\{\tilde{W}^T\dot{\tilde{W}}\right\} + e^T\dot{e} + tr\left\{\tilde{V}^T\dot{\tilde{V}}\right\} \tag{41}$$

The derivative of $V_1 = V(x)$ can be written as

$$\begin{aligned}\dot{V}_1 &= W^T(\nabla\sigma(x)\dot{x}) + (\alpha_1 + \alpha_2)x^T x \\ &\quad - \alpha_1 x^T(t - d_x)x(t - d_x) \\ &\quad - \alpha_2 x^T(t - d_u)x(t - d_u) + \nabla\epsilon(x)\dot{x}\end{aligned} \tag{42}$$

which by substituting the system dynamics (1), (42) becomes

$$\begin{aligned}\dot{V}_1 &= W^T\nabla\sigma(x)(Ax(t) + f(x(t - d_x) \\ &+ g_1\hat{u} + g_2\hat{u}(x(t - d_u))) + (\alpha_1 + \alpha_2)x^T x \\ &- \alpha_1 x^T(t - d_x)x(t - d_x) - \alpha_2 x^T(t - d_u)x(t - d_u) \\ &+ \nabla\epsilon(x)(Ax(t) + f(x(t - d_x) \\ &+ g_1\hat{u} + g_2\hat{u}(x(t - d_u)))\end{aligned} \tag{43}$$

Substituting the estimated control input (39) into (43) yields

$$\begin{aligned}\dot{V}_1 &= W^T\nabla\sigma(Ax(t) + f(x(t - d_x))) \\ &- g_1 W^T\nabla\sigma R^{-1}\hat{G}^T\nabla^T\sigma\hat{W} \\ &- g_2 W^T\nabla\sigma R^{-1}\hat{G}^T\nabla^T\sigma(x(t - d_u))\hat{W} \\ &- \alpha_1 x^T(t - d_x)x(t - d_x) - \alpha_2 x^T(t - d_u)x(t - d_u) \\ &+ (\alpha_1 + \alpha_2)x^T x + \nabla\epsilon(Ax(t) + f(x(t - d_x))) \\ &- g_1 W^T\nabla\epsilon R^{-1}\hat{G}^T\nabla^T\sigma\hat{W} \\ &- g_2 W^T\nabla\epsilon R^{-1}\hat{G}^T\nabla^T\sigma(x(t - d_u))\hat{W}\end{aligned} \tag{44}$$

Now, replacing the weight estimation error defined as $\hat{W} = W - \tilde{W}$ and $\tilde{G} = G - \hat{G}$ in (44), one gets

$$\begin{aligned}\dot{V}_1 &= W^T \nabla\sigma(Ax(t) + f(x(t-d_x)) \\ &- g_1 W^T \nabla\sigma R^{-1}(G - \tilde{G})^T \nabla^T \sigma(W - \tilde{W}) \\ &- g_2 W^T \nabla\sigma R^{-1}(G - \tilde{G})^T \nabla^T \sigma(x(t-d_u))(W - \tilde{W}) \\ &- \alpha_1 x^T(t-d_x)x(t-d_x) - \alpha_2 x^T(t-d_u)x(t-d_u) \\ &+ \nabla\epsilon(Ax(t) + f(x(t-d_x)) + (\alpha_1 + \alpha_2)x^T x \\ &- g_1 W^T \nabla\epsilon R^{-1}(G - \tilde{G})^T \nabla^T \sigma(W - \tilde{W}) \\ &- g_2 W^T \nabla\epsilon R^{-1}(G - \tilde{G})^T \nabla^T \sigma(x(t-d_u))(W - \tilde{W})\end{aligned} \tag{45}$$

which after some simplifications by separating the terms with respect to the actual NN weights and the terms with respect to the weight estimation error, (45) can be rewritten as

$$\begin{aligned}\dot{V}_1 &= W^T \nabla\sigma(Ax(t) + f(x(t-d_x)) \\ &- \alpha_1 x^T(t-d_x)x(t-d_x) - \alpha_2 x^T(t-d_u)x(t-d_u) \\ &+ \nabla\epsilon(Ax(t) + f(x(t-d_x)) + (\alpha_1 + \alpha_2)x^T x \\ &- g_1 W^T \nabla\sigma R^{-1} G^T \nabla^T \sigma W + g_1 W^T \nabla\sigma R^{-1} G^T \nabla^T \sigma \tilde{W} \\ &+ g_1 W^T \nabla\sigma R^{-1} \tilde{G}^T \nabla^T \sigma W - g_1 W^T \nabla\sigma R^{-1} \tilde{G}^T \nabla^T \sigma \tilde{W} \\ &- g_2 W^T \nabla\sigma R^{-1} G^T \nabla^T \sigma(x(t-d_u))W + g_2 W^T \nabla\sigma R^{-1} G^T \nabla^T \sigma(x(t-d_u))\tilde{W} \\ &+ g_2 W^T \nabla\sigma R^{-1} \tilde{G}^T \nabla^T \sigma(x(t-d_u))W - g_2 W^T \nabla\sigma R^{-1} \tilde{G}^T \nabla^T \sigma(x(t-d_u))\tilde{W} \\ &- g_1 W^T \nabla\epsilon R^{-1} G^T \nabla^T \sigma W + g_1 W^T \nabla\epsilon R^{-1} G^T \nabla^T \sigma \tilde{W} \\ &+ g_1 W^T \nabla\epsilon R^{-1} \tilde{G}^T \nabla^T \sigma W - g_1 W^T \nabla\epsilon R^{-1} \tilde{G}^T \nabla^T \sigma \tilde{W} \\ &- g_2 W^T \nabla\epsilon R^{-1} G^T \nabla^T \sigma(x(t-d_u))W + g_2 W^T \nabla\epsilon R^{-1} G^T \nabla^T \sigma(x(t-d_u))\tilde{W} \\ &+ g_2 W^T \nabla\epsilon R^{-1} \tilde{G}^T \nabla^T \sigma(x(t-d_u))W - g_2 W^T \nabla\epsilon R^{-1} \tilde{G}^T \nabla^T \sigma(x(t-d_u))\tilde{W}\end{aligned} \tag{46}$$

Using Fact 1 and assuming the nonlinear system with actual optimal control input is bounded, i.e. $W^T\nabla\sigma(Ax(t)+f(x(t-d_x))-g_1W^T\nabla\sigma R^{-1}G^T\nabla^T\sigma W-g_2W^T\nabla\sigma R^{-1}G^T\nabla^T\sigma(x(t-d_u))W \leqslant -\bar{k}\|x\|^2$. Moreover, the effect of the NN approximation error in the system dynamics is also assumed to be bounded above, i.e. $\nabla\epsilon(Ax(t) + f(x(t-d_x)) - g_1W^T\nabla\epsilon R^{-1}G^T\nabla^T\sigma W - g_2W^T\nabla\epsilon R^{-1}G^T\nabla^T\sigma(x(t-d_u))W \leqslant \bar{\epsilon}\|x\|^2$. Therefore, using the bounds and after some simplifications, (46) turns into

$$
\begin{aligned}
&\dot{V}_1 \leqslant -(\bar{k} + \bar{\epsilon} - \alpha_1 - \alpha_2)\|x\|^2 \\
&- \alpha_1\|x(t-d_x)\|^2 - \alpha_2\|x(t-d_u)\|^2 \\
&+ g_1W^T\nabla\sigma R^{-1}G^T\nabla^T\sigma\tilde{W} \\
&+ g_1W^T\nabla\sigma R^{-1}\tilde{G}^T\nabla^T\sigma W \\
&- g_1W^T\nabla\sigma R^{-1}\tilde{G}^T\nabla^T\sigma\tilde{W} \\
&+ g_2W^T\nabla\sigma R^{-1}G^T\nabla^T\sigma(x(t-d_u))\tilde{W} \\
&+ g_2W^T\nabla\sigma R^{-1}\tilde{G}^T\nabla^T\sigma(x(t-d_u))W \\
&- g_2W^T\nabla\sigma R^{-1}\tilde{G}^T\nabla^T\sigma(x(t-d_u))\tilde{W} \\
&+ g_1W^T\nabla\epsilon R^{-1}G^T\nabla^T\sigma\tilde{W} \\
&+ g_1W^T\nabla\epsilon R^{-1}\tilde{G}^T\nabla^T\sigma W \\
&- g_1W^T\nabla\epsilon R^{-1}\tilde{G}^T\nabla^T\sigma\tilde{W} \\
&+ g_2W^T\nabla\epsilon R^{-1}G^T\nabla^T\sigma(x(t-d_u))\tilde{W} \\
&+ g_2W^T\nabla\epsilon R^{-1}\tilde{G}^T\nabla^T\sigma(x(t-d_u))W \\
&- g_2W^T\nabla\epsilon R^{-1}\tilde{G}^T\nabla^T\sigma(x(t-d_u))\tilde{W}
\end{aligned} \tag{47}
$$

Next, assuming the coefficients of the weight estimation error are bounded, due to the boundedness of the actual NN weights and the nonlinear input functions g_1 and g_2, as $g_1W^T\nabla\sigma R^{-1}G^T\nabla^T\sigma + g_2W^T\nabla\sigma R^{-1}G^T\nabla^T\sigma(x(t-d_u)) \leqslant \bar{k}_1$ and $g_1W^T\nabla\epsilon R^{-1}G^T\nabla^T\sigma +$

$g_2 W^T \nabla\epsilon R^{-1} G^T \nabla^T \sigma(x(t-d_u)) \leqslant \bar{k}_2$, (47) can be written as

$$\begin{aligned}
\dot{V}_1 &\leqslant -(\bar{k} + \bar{\epsilon} - \alpha_1 - \alpha_2)\|x\|^2 \\
&- \alpha_1\|x(t-d_x)\|^2 - \alpha_2\|x(t-d_u)\|^2 \\
&+ (\bar{k}_1 + \bar{k}_2)\|\tilde{W}\| \\
&+ g_1 W^T \nabla\sigma R^{-1} \tilde{G}^T \nabla^T \sigma W \\
&- g_1 W^T \nabla\sigma R^{-1} \tilde{G}^T \nabla^T \sigma \tilde{W} \\
&+ g_2 W^T \nabla\sigma R^{-1} \tilde{G}^T \nabla^T \sigma(x(t-d_u))W \\
&- g_2 W^T \nabla\sigma R^{-1} \tilde{G}^T \nabla^T \sigma(x(t-d_u))\tilde{W} \\
&+ g_1 W^T \nabla\epsilon R^{-1} \tilde{G}^T \nabla^T \sigma W \\
&- g_1 W^T \nabla\epsilon R^{-1} \tilde{G}^T \nabla^T \sigma \tilde{W} \\
&+ g_2 W^T \nabla\epsilon R^{-1} \tilde{G}^T \nabla^T \sigma(x(t-d_u))W \\
&- g_2 W^T \nabla\epsilon R^{-1} \tilde{G}^T \nabla^T \sigma(x(t-d_u))\tilde{W}
\end{aligned} \tag{48}$$

Since $\tilde{G}$ is a function of the weight estimation error $\tilde{V}$, the following bounds are also considered for the known terms as $g_1 W^T \nabla\sigma R^{-1} \tilde{G}^T \nabla^T \sigma W + g_2 W^T \nabla\sigma R^{-1} \tilde{G}^T \nabla^T \sigma(x(t-d_u))W \leqslant \bar{k}_3\|\tilde{V}\|$ and $g_1 W^T \nabla\epsilon R^{-1} \tilde{G}^T \nabla^T \sigma W + g_2 W^T \nabla\epsilon R^{-1} \tilde{G}^T \nabla^T \sigma(x(t-d_u))W \leqslant \bar{k}_4\|\tilde{V}\|$. Then, (48) can be illustrated by

$$\begin{aligned}
\dot{V}_1 &\leqslant -(\bar{k} + \bar{\epsilon} - \alpha_1 - \alpha_2)\|x\|^2 \\
&- \alpha_1\|x(t-d_x)\|^2 - \alpha_2\|x(t-d_u)\|^2 \\
&+ (\bar{k}_1 + \bar{k}_2)\|\tilde{W}\| + (\bar{k}_3 + \bar{k}_4)\|\tilde{V}\| \\
&- g_1 W^T \nabla\sigma R^{-1} \tilde{G}^T \nabla^T \sigma \tilde{W} \\
&- g_2 W^T \nabla\sigma R^{-1} \tilde{G}^T \nabla^T \sigma(x(t-d_u))\tilde{W} \\
&- g_1 W^T \nabla\epsilon R^{-1} \tilde{G}^T \nabla^T \sigma \tilde{W} \\
&- g_2 W^T \nabla\epsilon R^{-1} \tilde{G}^T \nabla^T \sigma(x(t-d_u))\tilde{W}
\end{aligned} \tag{49}$$

The rest of the terms are the cross product terms of $\tilde{G}$ and $\tilde{W}$ which, by using the Young's inequality and adding the constant terms together as $\bar{k}_4\|\tilde{W}\|^2$ and $\bar{k}_5\|\tilde{V}\|^2$, results in

$$\begin{aligned}\dot{V}_1 \leqslant & -(\bar{k} + \bar{\epsilon} - \alpha_1 - \alpha_2)\|x\|^2 \\ & - \alpha_1\|x(t - d_x)\|^2 - \alpha_2\|x(t - d_u)\|^2 \\ & + (\bar{k}_1 + \bar{k}_2)\|\tilde{W}\| + (\bar{k}_3 + \bar{k}_4)\|\tilde{V}\| \\ & + \bar{k}_4\|\tilde{W}\|^2 + \bar{k}_5\|\tilde{V}\|^2\end{aligned} \tag{50}$$

The derivative of the second term in the overall Lyapunov candidate function, $\dot{V}_2 = tr\left\{\tilde{W}^T\dot{\tilde{W}}\right\}$ is given in the proof of Theorem 1 as

$$\dot{V}_2 \leqslant -\beta\left\|\tilde{W}\right\|^2 + \beta\epsilon_{max}\|\tilde{W}\| \tag{51}$$

The derivative of the third term in the overall Lyapunov candidate function, i.e. $\dot{V}_3 = e^T\dot{e}$, by considering the state estimation error (36), can be written as

$$\dot{V}_3 = e^T(-Ke + \tilde{V}\boldsymbol{\sigma}(\boldsymbol{X})\hat{\hat{u}} + \varepsilon_I(x)) \tag{52}$$

Since $\boldsymbol{\sigma}(\boldsymbol{X})$ is a function of the state vector and delayed state vector, and $\hat{\hat{u}}$ includes the cross product of $\tilde{G}$ and the critic NN weight $\tilde{W}$ and also using Assumption 3 $\|\varepsilon_I(x)\|^2 \leqslant \bar{b}_0\|x\|^2 + \bar{b}_1\|x(t-d_x)\|^2 + \bar{b}_3\|x(t-d_u)\|^2$, using the Young's inequality for the cross-product terms and the bounds defined for the NN identifier reconstruction error, gives the following simplified equation of (52) as

$$\begin{aligned}\dot{V}_3 \leqslant & -K\|e\|^2 + \bar{k}_6\|\tilde{V}\|^2 + \bar{k}_7\|\tilde{W}\|^2 + \bar{k}_8\|e\| \\ & + c_1\|x\|^2 + c_2\|x(t - d_x)\|^2 + \bar{c}_3\|x(t - d_u)\|^2\end{aligned} \tag{53}$$

Next, the derivative of the fourth term in the overall Lyapunov candidate function, i.e. $\dot{V}_4 = tr\{\tilde{V}^T\dot{\tilde{V}}\}$, by considering the weight estimation error (38) becomes

$$\dot{V}_4 = tr\{\tilde{V}^T\left(-\alpha_v\tilde{V} - \boldsymbol{\sigma}(\boldsymbol{X})\hat{\bar{u}}e^T + \alpha_v V\right)\} \tag{54}$$

which can be simplified as the cross product of the error terms as

$$\begin{aligned}\dot{V}_4 &\leqslant -\alpha_v\|\tilde{V}\|^2 + \alpha_v c_3\|\tilde{V}\| + c_4\|\tilde{W}\|^2 + c_5\|e\|^2 \\ &+ c_6\|x\|^2 + c_7\|x(t-d_x)\|^2\end{aligned} \tag{55}$$

Collecting the derivatives of four corresponding terms in the overall Lyapunov candidate function under the PE condition, yields

$$\begin{aligned}\dot{L} &\leqslant -(\bar{k} + \bar{\epsilon} - \alpha_1 - \alpha_2)\|x\|^2 \\ &- \alpha_1\|x(t-d_x)\|^2 - \alpha_2\|x(t-d_u)\|^2 \\ &+ (\bar{k}_1 + \bar{k}_2)\|\tilde{W}\| + (\bar{k}_3 + \bar{k}_4)\|\tilde{V}\| \\ &+ \bar{k}_4\|\tilde{W}\|^2 + \bar{k}_5\|\tilde{V}\|^2 \\ &- \beta\left\|\tilde{W}\right\|^2 + \beta\epsilon_{max}\|\tilde{W}\| \\ &- K\|e\|^2 + \bar{k}_6\|\tilde{V}\|^2 + \bar{k}_7\|\tilde{W}\|^2 + \bar{k}_8\|e\| \\ &+ c_1\|x\|^2 + c_2\|x(t-d_x)\|^2 + \bar{c}_3\|x(t-d_u)\|^2 \\ &- \alpha_v\|\tilde{V}\|^2 + \alpha_v c_3\|\tilde{V}\| + c_4\|\tilde{W}\|^2 + c_5\|e\|^2 \\ &+ c_6\|x\|^2 + c_7\|x(t-d_x)\|^2\end{aligned} \tag{56}$$

which by defining $\theta_1 = \bar{k} + \bar{\epsilon} + c_1 + c_6 - \alpha_1 - \alpha_2$, $\theta_2 = \alpha_1 - c_2 - c_7$, $\theta_3 = \beta - \bar{k}_7 - \bar{k}_4 - c_4$, $\theta_4 = \bar{k}_1 + \bar{k}_2 + \beta\epsilon_{max}$, $\theta_5 = \alpha_v - \bar{k}_5 - \bar{k}_6$, $\theta_6 = \alpha_v c_3 + \bar{k}_3 + \bar{k}_4$ and $\theta_7 = K - c_5$ becomes

$$\begin{aligned}\dot{L} &\leqslant -\theta_1\|x\|^2 - \theta_2\|x(t-d_x)\|^2 \\ &- \theta_3\left\|\tilde{W}\right\|^2 + \theta_4\|\tilde{W}\| - \theta_5\|\tilde{V}\|^2 + \theta_6\|\tilde{V}\| \\ &- \theta_7\|e\|^2 + \bar{k}_8\|e\| - (\alpha_2 - \bar{c}_3)\|x(t-d_u)\|^2\end{aligned} \tag{57}$$

Then, after some manipulation, (57) gives

$$\begin{aligned}\dot{L} &\leqslant -\theta_1\|x\|^2 - \theta_2\|x(t-d_x)\|^2 \\ &- (\alpha_2 - \bar{c}_3)\|x(t-d_u)\|^2 \\ &- \theta_3\left(\left(\|\tilde{W}\| - \frac{\theta_4}{2\theta_3}\right)^2 - \frac{\theta_4^2}{4\theta_3^2}\right) \\ &- \theta_5\left(\left(\|\tilde{V}\| - \frac{\theta_6}{2\theta_5}\right)^2 - \frac{\theta_6^2}{4\theta_5^2}\right) \\ &- \theta_7\left(\left(\|e\| - \frac{\bar{k}_8}{2\theta_7}\right)^2 - \frac{\bar{k}_8^2}{4\theta_7^2}\right)\end{aligned} \tag{58}$$

which by summing up the constant terms as $\Gamma = \frac{\theta_4^2}{4\theta_3} + \frac{\theta_6^2}{4\theta_5} + \frac{\bar{k}_8^2}{4\theta_7}$, becomes

$$\begin{aligned}\dot{L} &\leqslant -\theta_1\|x\|^2 - \theta_2\|x(t-d_x)\|^2 \\ &- \theta_3\left(\|\tilde{W}\| - \frac{\theta_4}{2\theta_3}\right)^2 - \theta_5\left(\|\tilde{V}\| - \frac{\theta_6}{2\theta_5}\right)^2 \\ &- \theta_7\left(\|e\| - \frac{\bar{k}_8}{2\theta_7}\right)^2 - (\alpha_2 - \bar{c}_3)\|x(t-d_u)\|^2 + \Gamma\end{aligned} \tag{59}$$

As exemplified in (59), the overall Lyapunov candidate function L becomes less than zero, i.e. $\dot{L} < 0$ if $\theta_1 > 0$, $\theta_2 > 0$, $\theta_3 > 0$, $\theta_4 > 0$, $\theta_5 > 0$, $\theta_6 > 0$, $\theta_7 > 0$, and $\alpha_2 > \bar{c}_3$ provided one of the following conditions hold

$$\begin{cases} \|x\| > \sqrt{\dfrac{\Gamma}{\theta_1}} \\ \|x(t-d_x)\| > \sqrt{\dfrac{\Gamma}{\theta_2}} \\ \|x(t-d_u)\| > \sqrt{\dfrac{\Gamma}{(\alpha_2 - \bar{c}_3)}} \\ \|\tilde{W}\| > \sqrt{\dfrac{\Gamma}{\theta_3}} + \dfrac{\theta_4}{2\theta_3} \\ \|\tilde{V}\| > \sqrt{\dfrac{\Gamma}{\theta_5}} + \dfrac{\theta_6}{2\theta_5} \\ \|e\| > \sqrt{\dfrac{\Gamma}{\theta_7}} + \dfrac{\bar{k}_8}{2\theta_7} \end{cases}$$

This confirms the boundedness of the overall closed-loop system, including the NN weights. To show the boundedness of the optimal control policy, the control input error can be defined as $\tilde{u} = u - \hat{u}$. Substituting the actual control input (7) and the estimated control input (19), one has $\tilde{u} = -R^{-1}G^T\nabla^T\sigma(x)W - R^{-1}G^T(x(t))\nabla\epsilon(x) + R^{-1}\hat{G}^T\nabla^T\sigma(x)\hat{W}$. Then, considering $\tilde{W} = W - \hat{W}$ and $\hat{G} = G - \tilde{G}$ with $\tilde{G}$ as the control input coefficient matrix error, one has $\tilde{u} = -R^{-1}G^T\nabla^T\sigma(x)\tilde{W} - R^{-1}\tilde{G}^T\nabla^T\sigma(x)W + R^{-1}\tilde{G}^T\nabla^T\sigma(x)\tilde{W} - R^{-1}G^T(x(t))\nabla\epsilon(x)$. The last term in the control input error $R^{-1}G^T(x(t))\nabla\epsilon(x)$ is bounded as the state vector is bounded and G is also bounded due to the boundedness of the nonlinear input matrices g_1 and g_2 and because of the increasing the number of neurons results in the convergence of $\nabla\epsilon(x)$ to zero. The input coefficient matrix $\hat{G}$ is obtained from the the NN weights of the identifier, therefore, $\tilde{G}$ is a function of the NN of the NN weight estimation error of the identifier, i.e. $\tilde{V}$. As a consequence, it is shown in the first part of Theorem 3 that $\tilde{W}$, $\tilde{V}$, system state and the NN identification error are. However, the control input error becomes bounded which shows that the estimated control policy converges to the optimal control policy. ■

Remark 11. For the boundedness of the overall closed-loop system, the Lyapunov function candidate also includes quadratic terms of the NN identifier, such as state estimation and second NN weight estimation errors. To ensure the Lyapunov function is less than zero, the history of the states have to be selected in a specific bound. Moreover, bounds are functions of the design parameters given in the definition of the value function.

The simulation result is presented in the next section to ensure the effectiveness of the proposed approach.

4. SIMULATION RESULTS

In this section, a simulation example is provided for a continuous time nonlinear system with state and input delays which substantiate the theoretical claims for both cases- when the control input coefficient matrix G is known and when it is unknown.

Example 1. Consider the following nonlinear continuous time-delay system with state ad input delay as

$$\begin{aligned} \dot{x} &= \begin{bmatrix} 0.1 & 1 \\ 0 & -1 \end{bmatrix} x(t) + \begin{bmatrix} -\frac{x_1(t-d_x)}{\sqrt{1+x_1^2(t-d_x)}} \\ \frac{x_2(t-d_x)}{\sqrt{1+x_2^2(t-d_x)}} \end{bmatrix} \\ &+ \begin{bmatrix} -\frac{1}{\sqrt{1+x_1^2(t-d_x)}} \\ 0 \end{bmatrix} u(t) + \begin{bmatrix} 0 \\ \frac{1}{\sqrt{1+x_1^2(t-d_x)}} \end{bmatrix} u(t-d_u) \end{aligned} \tag{60}$$

The values for the delays are considered as $d_x = 0.001$ and $d_u = 0.001$. The history functions are assumed to be $x(\theta) = 1, \theta \in [-d_x, 0]$ and $u(\theta) = 0, \theta \in [-d_u, 0]$. The parameter of the quadratic cost function are designed as $Q = 10I_{2\times2}$ and $R = 1$.

4.1. CONTROL COEFFICIENT G MATRIX KNOWN

In this subsection, it is assumed that the coefficient matrix G is known. The log-sigmoid activation function is utilized for the NN which is a vector with the dimension of 2×1, as it is a function of the system states and NN weights are initialized as $\hat{W} = [0.50.5]^T$. The value of coefficient of the integral term in the value function is defined as $\alpha = 1.9$.

To compute the gain matrix G as defined in (7), the matrix A_0 and the input delay d_u are needed. The delay value is known and A_0 can be obtained from Assumption 1 by using a nominal model of the system. However, since the internal dynamic of the given system is already linear, i.e. $\begin{bmatrix} 0.1 & 1 \\ 0 & -1 \end{bmatrix}$, for the simulation, the matrix A_0 is considered to be equal to A which is an unstable open-loop system.

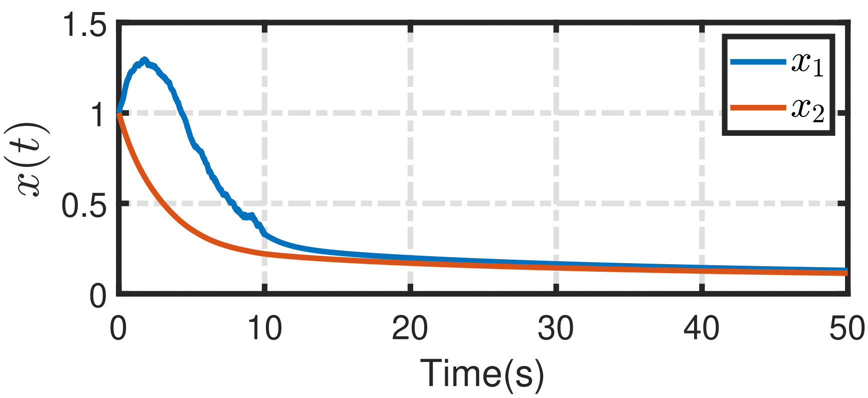


Figure 3. State vector of the nonlinear system (60).

For the simulation, a random noise is added to the control input for $t < 10$. The performance of the proposed controller is shown in Figures 3, 4 and 5. It can be seen from Figure 5 that the estimated neural network weight converges, which shows the convergence of the approximated critic function. This justifies the theoretical claims in Theorem 1. Moreover, it can be observed in Figure 3 that the states of the system become bounded, which substantiate the result of Theorem 2.

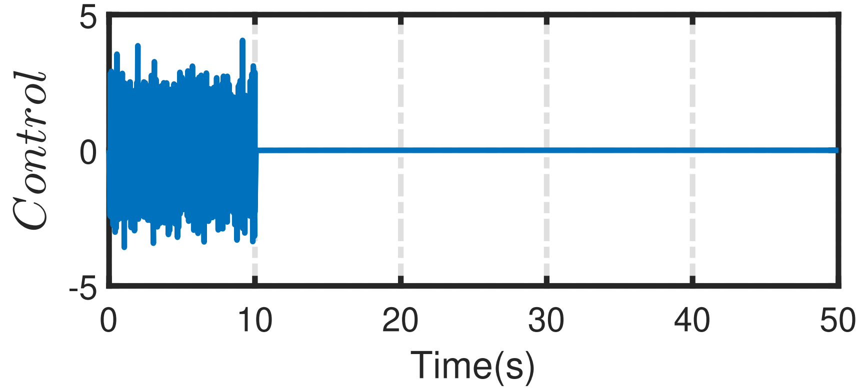


Figure 4. Control input under the PE condition.

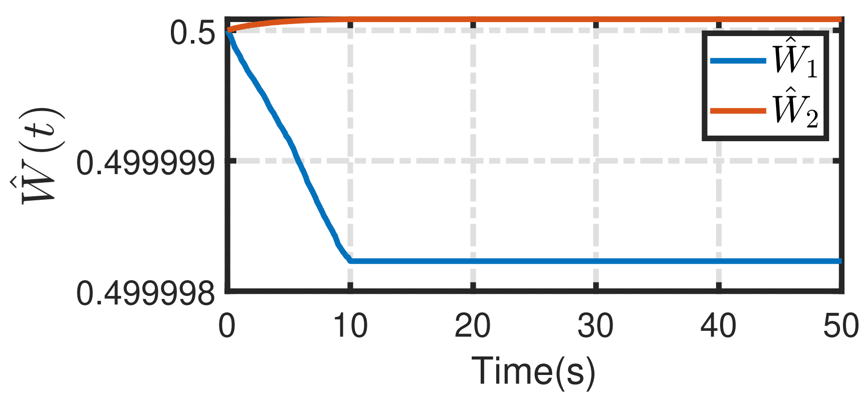


Figure 5. Estimated NN weights.

4.2. CONTROL COEFFICIENT G MATRIX UNKNOWN

In this subsection, it is assumed that the matrix G is unknown. The identifier (34) is used to estimate the matrices g_1 and g_2 and, then, $\hat{G}$ is utilized in the estimated control input (39). It is assumed that the nominal value of the internal matrix A_0 is available. The log-sigmoid activation function is utilized for both the critic NN, i.e. $\hat{W}$ and also the estimator NN, i.e. $\hat{V}$. The dimension of the critic NN is 2×1, as it is a function of the system states and its weights are initialized as $\hat{W} = [0.5, 0.5]^T$. The value of coefficient of the integral term in the value function is defined as $\alpha = 1.9$. The dimension of the estimator NN is 8×2, as it is a function of the system states and its weights are initialized as $\hat{V} = 0.01ones(8, 2)$. The matrix A_0 is considered to be equal to A which is an unstable open-loop system. The identifier parameters are considered as $K = 10$ and $\alpha_v = 0.2$ and the identifier is initialized at the system state, i.e. $\hat{x}_0 = x_0$.

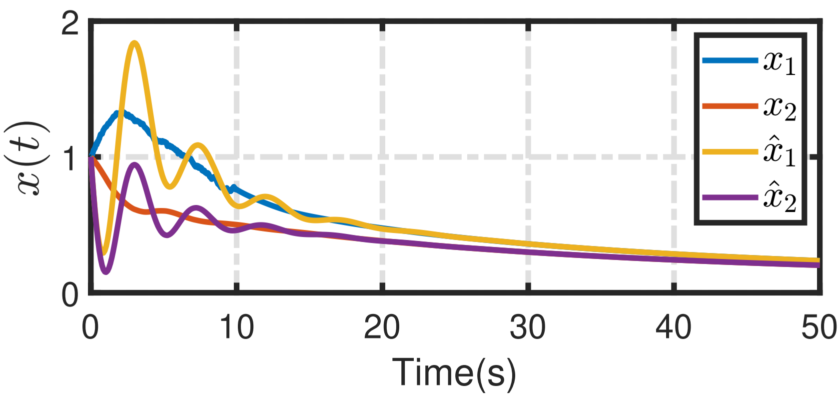


Figure 6. The state vector of the nonlinear system (60) and the NN identifier state (34).

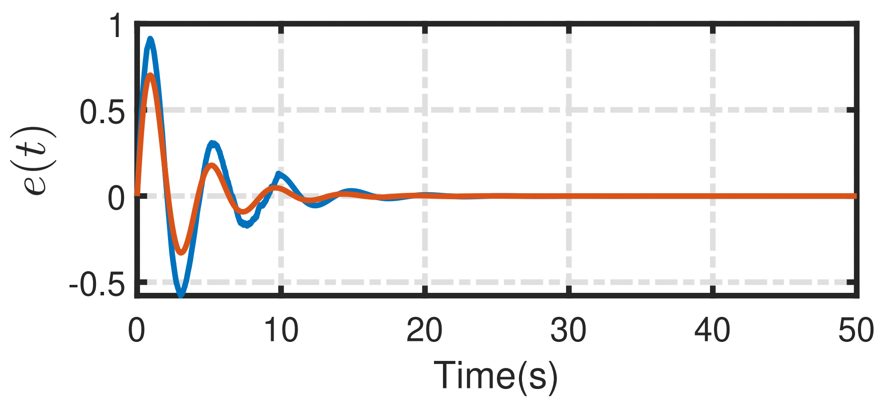


Figure 7. The NN identifier estimation error.

The performance of the controller is shown in Figures 6, 7, and 8. To satisfy the PE condition, a random noise is added to the control input when $t < 10$. It can be observed from Figure 6 that the system state vector and the identifier state become bounded eventually. Moreover, the performance of the identifier is shown in Figure 7. It can be seen that the estimation error converges eventually. The control input signal is shown in Figure 8 which appears to be near zero. All the results provided in the simulation, confirm the results of Theorems 1 and 3.

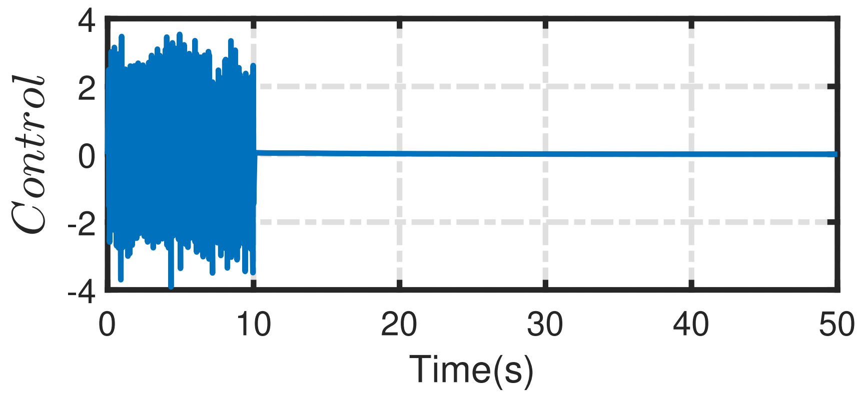


Figure 8. Control input under the PE condition.

5. CONCLUSION

A NN based optimal adaptive regulation for a class of nonlinear continuous-time systems with state and input delays and under partial and complete uncertain dynamics is studied. The linearized internal dynamics is necessary to relate the derivative of the delayed input vector with respect to the current input. The known control coefficient matrix helps in simplifying the problem, whereas, one can relax the need for control coefficient matrix through the addition of an actor NN or an identifier as given in the paper for time-delay systems. The NN identifier weights, based on the product of augmented input and state estimation error, approximated the control coefficient matrix, which is employed in the computation of estimated control policy. The NN based actor-critic framework with NN identifier estimated the optimal control input and cost function. The ITDE using IRL tunes the actor NN weights. The boundedness of the state vector and critic NN weights are shown through Lyapunov analysis. Finally, it is proven that the estimated control policy is bounded close to the optimal control policy. Simulation results are provided to illustrate the effectiveness of the proposed approach.

REFERENCES

[1] J. P. Hespanha, P. Naghshtabrizi, and Y. Xu. A survey of recent results in networked control systems. Proceedings of the IEEE, 95(1):138–162, Jan 2007.

[2] Keqin Gu, Jie Chen, and Vladimir L Kharitonov. Stability of time-delay systems. Springer Science & Business Media, 2003.

[3] Emilia Fridman. Introduction to Time-delay Systems: Analysis and Control. Springer, 2014.

[4] Jean-Pierre Richard. Time-delay systems: an overview of some recent advances and open problems. Automatica, 39(10):1667 – 1694, 2003.

[5] Iasson Karafyllis, Michael Malisoff, Frederic Mazenc, and Pierdomenico Pepe. Recent results on nonlinear delay control systems. Springer International Publishing Switzerland, 2016.

[6] Y. Zhu, M. Krstic, and H. Su. Adaptive output feedback control for uncertain linear time-delay systems. IEEE Transactions on Automatic Control, 62(2):545–560, Feb 2017.

[7] A. Ponomarev. Reduction-based robustness analysis of linear predictor feedback for distributed input delays. IEEE Transactions on Automatic Control, 61(2):468–472, Feb 2016.

[8] Z. Artstein. Linear systems with delayed controls: A reduction. IEEE Transactions on Automatic Control, 27(4):869–879, August 1982.

[9] X. Cai, N. Bekiaris-Liberis, and M. Krstic. Input-to-state stability and inverse optimality of linear time-varying-delay predictor feedbacks. IEEE Transactions on Automatic Control, 63(1):233–240, Jan 2018.

[10] V. L. Kharitonov, S. . Niculescu, J. Moreno, and W. Michiels. Static output feedback stabilization: necessary conditions for multiple delay controllers. IEEE Transactions on Automatic Control, 50(1):82–86, Jan 2005.

[11] D. Bresch-Pietri and M. Krstic. Delay-adaptive predictor feedback for systems with unknown long actuator delay. IEEE Transactions on Automatic Control, 55(9):2106–2112, Sep. 2010.

[12] Miroslav Krstic. Delay compensation for nonlinear, adaptive, and PDE systems. Springer, 2009.

[13] Yang Zhu, Miroslav Krstic, and Hongye Su. Adaptive global stabilization of uncertain multi-input linear time-delay systems by pde full-state feedback. Automatica, 96:270 – 279, 2018.

[14] Kim-Doang Nguyen. An adaptive controller for uncertain nonlinear systems with multiple time delays. Automatica, 117:108976, 2020.

[15] Delphine Bresch-Pietri, Jonathan Chauvin, and Nicolas Petit. Adaptive control scheme for uncertain time-delay systems. Automatica, 48(8):1536 – 1552, 2012.

[16] H. Zhang, Y. Liu, G. Xiao, and H. Jiang. Data-based adaptive dynamic programming for a class of discrete-time systems with multiple delays. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 50(2):432–441, 2020.

[17] Y. Liu, H. Zhang, R. Yu, and Z. Xing. H_∞ tracking control of discrete-time system with delays via data-based adaptive dynamic programming. IEEE Transactions on Systems, Man, and Cybernetics: Systems, pages 1–8, 2019.

[18] D. Ross and I. Flügge-Lotz. An optimal control problem for systems with differential-difference equation dynamics. SIAM Journal on Control, 7(4):609–623, 1969.

[19] M. Basin and J. Rodriguez-Gonzalez. Optimal control for linear systems with multiple time delays in control input. IEEE Transactions on Automatic Control, 51(1):91–97, Jan 2006.

[20] D. Maity, M. H. Mamduhi, S. Hirche, K. H. Johansson, and J. S. Baras. Optimal lqg control under delay-dependent costly information. IEEE Control Systems Letters, 3(1):102–107, Jan 2019.

[21] R. Moghadam and S. Jagannathan, "Optimal control of linear continuous-time systems in the presence of state and input delays with application to a chemical reactor," in 2020 American Control Conference (ACC), pp. 999–1004, 2020.

[22] R. Moghadam and S. Jagannathan, "Approximate optimal adaptive control of partially unknown linear continuous-time systems with state delay," in 2019 IEEE 58th Conference on Decision and Control (CDC), pp. 1985–1990, 2019.

[23] F. Cacace, F. Conte, A. Germani, and G. Palombo. Optimal control of linear systems with large and variable input delays. Systems & Control Letters, 89:1 – 7, 2016.

[24] P. Ioannou and B. Fidan, Adaptive control tutorial, vol. 11. Siam, 2006.

[25] Richard S Sutton and Andrew G Barto. Reinforcement learning: An introduction. MIT press, 2018.

[26] F. L. Lewis and D. Vrabie. Reinforcement learning and adaptive dynamic programming for feedback control. IEEE Circuits and Systems Magazine, 9(3):32–50, 2009.

[27] V. Narayanan and S. Jagannathan. Event-triggered distributed control of nonlinear interconnected systems using online reinforcement learning with exploration. IEEE Transactions on Cybernetics, 48(9):2510–2519, Sep. 2018.

[28] F. L. Lewis, D. Vrabie, and V. L. Syrmos, Optimal Control. John Wiley & Sons, 2012.

[29] J. Sarangapani, Neural Network Control of Nonlinear Discrete-time Systems. CRC press, 2006.

[30] D. E. Kirk, Optimal Control Theory: An Introduction. Courier Corporation, 2004.

[31] D. P. Bertsekas, "Value and policy iterations in optimal control and adaptive dynamic programming," IEEE Transactions on Neural Networks and Learning Systems, vol. 28, no. 3, pp. 500–509, 2015.

[32] F. Al-Tamimi and M. Abu-Khalaf, "Discrete-time nonlinear hjb solution using approximate dynamic programming: Convergence proof," IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), vol. 38, no. 4, pp. 943–949, 2008.

[33] G. Xiao, H. Zhang, and Y. Luo, "Online optimal control of unknown discrete-time nonlinear systems by using time-based adaptive dynamic programming," Neurocomputing, vol. 165, pp. 163–170, 2015.

[34] T. Dierks and S. Jagannathan, "Online optimal control of affine nonlinear discrete-time systems 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, 2012.

[35] E. N. Sanchez and F. Ornelas-Tellez, Discrete-time inverse optimal control for nonlinear systems. CRC Press, 2017.

[36] F. L. Lewis and D. Liu, Reinforcement Learning and Approximate Dynamic Programming for Feedback Control, vol. 17. John Wiley & Sons, 2013.

[37] O. Santos, S. Mondié, and V.L. Kharitonov. Linear quadratic suboptimal control for time delays systems. International Journal of Control, 82(1):147–154, 2009.

[38] A. Manitius and A. Olbrot. Finite spectrum assignment problem for systems with delays. IEEE Transactions on Automatic Control, 24(4):541–552, 1979.

[39] J. Zhu, Y. Hou, and T. Li, "Optimal control of nonlinear systems with time delays: An online adp perspective," IEEE Access, vol. 7, pp. 145574–145581, 2019.

[40] Y. Niu, D. W C Ho, and J. Lam, "Robust integral sliding mode control for uncertain stochastic systems with time-varying delay," Automatica, vol. 41, no. 5, pp. 873 – 880, 2005.

[41] S. S. Ge, Fan Hong, and Tong Heng Lee, "Adaptive neural network control of nonlinear systems with unknown time delays," IEEE Transactions on Automatic Control, vol. 48, no. 11, pp. 2004–2010, 2003.

[42] P. Pepe, "Discrete-time systems with constrained time delays and delay-dependent lyapunov functions," IEEE Transactions on Automatic Control, vol. 65, no. 4, pp. 1724–1730, 2020.

[43] S. Battilotti, “Continuous-time and sampled-data stabilizers for nonlinear systems with input and measurement delays,” IEEE Transactions on Automatic Control, vol. 65, no. 4, pp. 1568–1583, 2020.

[44] Gilbert Strang. Introduction to Linear Algebra. Cambridge Press, 2016.

[45] Frank L Lewis, Draguna Vrabie, and Vassilis L Syrmos. Optimal Control. John Wiley & Sons, 2012.

[46] E. D. Sontag and Yuan Wang. New characterizations of input-to-state stability. IEEE Transactions on Automatic Control, 41(9):1283–1294, 1996.

[47] P. Pepe and Z.-P. Jiang. A lyapunov–krasovskii methodology for iss and iiss of time-delay systems. Systems & Control Letters, 55(12):1006 – 1014, 2006.

[48] Hassan K Khalil. Nonlinear control. Pearson New York, 2015.

[49] G. Cybenko, “Approximation by superpositions of a sigmoidal function,” Mathematics of control, signals and systems, vol. 2, no. 4, pp. 303–314,1989.

[50] F. L. Lewis and J. Campos and R. Selmic,“Neuro-Fuzzy Control of Industrial Systems with Actuator Nonlinearities,” Society for Industrial and Applied Mathematics, 2002.

SECTION

2. CONCLUSIONS AND FUTURE WORK

In this dissertation, a suite of optimal adaptive control techniques are utilized to develop finite dimensional model-free optimal control of linear and nonlinear time-delay systems. Initially the linear dynamics of the systems are considered known, whereas this assumption is relaxed by taking the dynamics as the linear in the unknown parameters (LIP). Traditional adaptive techniques are embedded with optimal control to derive optimal adaptive control methods for such linear uncertain systems using state and output feedback. The possibility of minimizing the transmission of the state vector to the controller is accomplished along the way by deriving an event trigger condition. Finally, the nonlinear dynamics are considered uncertain and neural networks are employed to approximate them. The neural network based optimal adaptive controller relaxes the need for LIP assumption while needing the measurement of state vector.

2.1. CONCLUSIONS

First, the optimal control of linear continuous-time systems with state and input delays by utilizing a quadratic cost function over infinite horizon is addressed using both state and output feedback. The value function is defined in terms of the system state and system output and, then, using the stationarity conditions, the optimal controls are obtained. For both cases, the Lyapunov-Karkovskii value function and Bellman type equation facilitated the optimal control policy through delay algebraic Ricatti equations (DARE) without using state transformation or need to solve PDE or ODE. The DARE is a function of state and input delays that are considered to be known.

Next, the optimal adaptive regulation of unknown linear continuous-time systems with state and input delays by using output feedback is attempted under a mild assumption that the system matrices are uncertain but their bounds are known. First, an adaptive identifier is utilized with a Bellman type equation to minimize the quadratic cost function. By using the estimated DARE and Bellman type equation, the estimated optimal control input is shown to approach the actual optimal control input with a bounded error for the case of uncertain dynamics with output feedback. The persistent exciting noise ensured the convergence of the estimated parameters and provided the initial estimated parameters in a compact set.

Subsequently, the optimal adaptive control of partially unknown linear continuous time systems with state-delay by using IRL is addressed. It was demonstrated that the optimal control policy renders asymptotic stability of the closed-loop system provided that the linear time delayed system is controllable and observable. The event-triggered approach is used to relax the need for continuous availability of the system state. Next, the IRL approach using actor-critic structure estimates the optimal control input without requiring any information about drift dynamics. However, the input matrix is still needed to find the control policy. Event sampled and hybrid approaches are introduced to update the value function estimation. The Lyapunov theory is employed to show boundedness of the overall closed-loop system. Simulation examples verify the effectiveness of the proposed methods.

Finally, an optimal adaptive control (OAC) for uncertain nonlinear continuous time systems with input and state delays by using IRL is studied. First, a quadratic cost function over infinite time horizon is considered and a value function is defined by considering the delayed state of the system. Subsequently, an actor-critic framework is utilized using the IRL approach to relax the need for system dynamics for control policy generation. The critic network is approximated through a neural network and temporal difference error, derived from the difference between the actual and approximated value function, is utilized to find the update law for the weights of the critic function. It is assumed that the control coefficient matrix is known, then, this assumption is relaxed by introducing a neural network based identifier. Lyapunov theory is employed to demonstrate the boundedness of the overall closed-loop system.

2.2. FUTURE WORK

As part of the future work, the controller can be designed for time-delay systems with input and state delays in the presence the uncertain system dynamics as well as uncertain delays using the learning approaches. Though the state feedback is considered for nonlinear time-delay systems, output feedback is more preferred. Moreover, the effect of adversaries to the time delay systems such as disturbances and attacks can also be considered in the design of the controller to make the overall system resilient. This can be challenging because the attack can be a source of the delay to degrade the overall network performance.

REFERENCE

[1] Keqin Gu, Jie Chen, and Vladimir L Kharitonov. Stability of time-delay systems. Springer Science & Business Media, 2003.

[2] Emilia Fridman. Introduction to Time-delay Systems: Analysis and Control. Springer, 2014.

[3] Jean-Pierre Richard. Time-delay systems: an overview of some recent advances and open problems. Automatica, 39(10):1667 – 1694, 2003.

[4] Iasson Karafyllis, Michael Malisoff, Frederic Mazenc, and Pierdomenico Pepe. Recent results on nonlinear delay control systems. Springer International Publishing Switzerland, 2016.

[5] Y. Zhu, M. Krstic, and H. Su. Adaptive output feedback control for uncertain linear time-delay systems. IEEE Transactions on Automatic Control, 62(2):545–560, Feb 2017.

[6] A. Ponomarev. Reduction-based robustness analysis of linear predictor feedback for distributed input delays. IEEE Transactions on Automatic Control, 61(2):468–472, Feb 2016.

[7] Z. Artstein. Linear systems with delayed controls: A reduction. IEEE Transactions on Automatic Control, 27(4):869–879, August 1982.

[8] X. Cai, N. Bekiaris-Liberis, and M. Krstic. Input-to-state stability and inverse optimality of linear time-varying-delay predictor feedbacks. IEEE Transactions on Automatic Control, 63(1):233–240, Jan 2018.

[9] V. L. Kharitonov, S. . Niculescu, J. Moreno, and W. Michiels. Static output feedback stabilization: necessary conditions for multiple delay controllers. IEEE Transactions on Automatic Control, 50(1):82–86, Jan 2005.

[10] D. Bresch-Pietri and M. Krstic. Delay-adaptive predictor feedback for systems with unknown long actuator delay. IEEE Transactions on Automatic Control, 55(9):2106–2112, Sep. 2010.

[11] Miroslav Krstic. Delay compensation for nonlinear, adaptive, and PDE systems. Springer, 2009.

[12] S. Su, Y. Wei, and Z. Lin. Stabilization of discrete-time linear systems with an unknown time-varying delay by switched low-gain feedback. IEEE Transactions on Automatic Control, 64(5):2069–2076, May 2019.

[13] D. Ross and I. Flügge-Lotz. An optimal control problem for systems with differential-difference equation dynamics. SIAM Journal on Control, 7(4):609–623, 1969.

[14] Michael Basin, Jesus Rodriguez-Gonzalez, and Leonid Fridman. Optimal and robust control for linear state-delay systems. Journal of the Franklin Institute, 344(6):830 – 845, 2007.

[15] M. Basin and J. Rodriguez-Gonzalez. Optimal control for linear systems with multiple time delays in control input. IEEE Transactions on Automatic Control, 51(1):91–97, Jan 2006.

[16] Yusheng Zhou and Zaihua Wang. A robust optimal trajectory tracking control for systems with an input delay. Journal of the Franklin Institute, 353(12):2627 – 2649, 2016.

[17] Guo-Ping Cai, Jin-Zhi Huang, and Simon X Yang. An optimal control method for linear systems with time delay. Computers & Structures, 81(15):1539 – 1546, 2003.

[18] O. Santos, S. Mondié, and V.L. Kharitonov. Linear quadratic suboptimal control for time delays systems. International Journal of Control, 82(1):147–154, 2009.

[19] Michael Basin and Jesus Rodriguez-Gonzalez. A closed-form optimal control for linear systems with equal state and input delays. Automatica, 41(5):915 – 920, 2005.

[20] Y. Fiagbedzi and A. Pearson. Feedback stabilization of linear autonomous time lag systems. IEEE Transactions on Automatic Control, 31(9):847–855, Sep. 1986.

[21] Shoulin Hao, Tao Liu, and Bin Zhou. Output feedback anti-disturbance control of input-delayed systems with time-varying uncertainties. Automatica, 104:8 – 16, 2019.

[22] Filippo Cacace and Alfredo Germani. Output feedback control of linear systems with input, state and output delays by chains of predictors. Automatica, 85:455 – 461, 2017.

[23] Le Van Hien and Mai Thi Hong. An optimization approach to static output-feedback control of lti positive systems with delayed measurements. Journal of the Franklin Institute, 2019.

[24] Saeed Ahmed, Frédéric Mazenc, and Hitay Özbay. Dynamic output feedback stabilization of switched linear systems with delay via a trajectory based approach. Automatica, 93:92 – 97, 2018.

[25] Chi-Chang Lin, Chang-Ching Chang, and Huang-Lin Chen. Optimal h_∞ output feedback control systems with time delay. Journal of Engineering Mechanics, 132(10):1096–1105, 2006.

[26] Se Young Yoon and Zongli Lin. Robust output regulation of linear time-delay systems: A state predictor approach. International Journal of Robust and Nonlinear Control, 26(8):1686–1704, 2016.

[27] V. Narayanan and S. Jagannathan. Event-triggered distributed control of nonlinear interconnected systems using online reinforcement learning with exploration. IEEE Transactions on Cybernetics, 48(9):2510–2519, Sep. 2018.

[28] Petros Ioannou and Baris Fidan. Adaptive control tutorial, volume 11. Siam, 2006.

[29] F. L. Lewis and D. Vrabie. Reinforcement learning and adaptive dynamic programming for feedback control. IEEE Circuits and Systems Magazine, 9(3):32–50, 2009.

[30] W. Gao and Z. Jiang. Adaptive optimal output regulation of time-delay systems via measurement feedback. IEEE Transactions on Neural Networks and Learning Systems, 30(3):938–945, March 2019.

[31] Y. Orlov, I. V. Kolmanovsky, and O. Gomez. Adaptive identification of linear time-delay systems: From theory toward application to engine transient fuel identification. International Journal of Adaptive Control and Signal Processing, 23(2):150–165, 2009.

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

[33] D. P. Borgers, V. S. Dolk, and W. P. M. H. Heemels. Riccati-based design of event-triggered controllers for linear systems with delays. IEEE Transactions on Automatic Control, 63(1):174–188, Jan 2018.

[34] 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, 58(2):422–434, Feb 2013.

[35] V. Narayanan and S. Jagannathan. Event-triggered distributed approximate optimal state and output control of affine nonlinear interconnected systems. IEEE Transactions on Neural Networks and Learning Systems, 29(7):2846–2856, July 2018.

[36] I. Karafyllis and M. Krstic. Adaptive certainty-equivalence control with regulation-triggered finite-time least-squares identification. IEEE Transactions on Automatic Control, 63(10):3261–3275, Oct 2018.

VITA

Rohollah Moghadam received his Bachelor of Science (B.S.) in Electrical Engineering form Azad University of Tehran, Tehran, Iran, in December 2001, and he earned a Master of Science (M.Sc.) in Electrical Engineering from Shahrood University of Technology, Shahrood, Iran in March 2007. He was a Ph.D. student of Amirkabir University of Technology, Tehran, Iran, in 2013 and transferred his credits. He was a visiting scholar at the University of Texas at Arlington Research Institute, Fort Worth, Texas, USA in 2016 and with the Missouri University of Science and Technology, Rolla, Missouri in 2017. He received a Ph.D. in Electrical Engineering from Missouri University of Science and Technology, Rolla, Missouri in December 2020.