A Hybrid Time-Scaling Transformation for Time-Delay Optimal Control Problems

Changjun Yu · Qun Lin · Ryan Loxton · Kok Lay Teo · Guoqiang Wang

Abstract In this paper, we consider a class of nonlinear time-delay optimal control problems with canonical equality and inequality constraints. We propose a new computational approach, which combines the control parameterization technique with a hybrid time-scaling strategy, for solving this class of optimal control problems. The proposed approach involves approximating the control variables by piecewise constant functions, whose heights and switching times are decision variables to be optimized. Then, the resulting problem with varying switching times is transformed, via a new hybrid time-scaling strategy, into an equivalent problem with fixed switching times, which is much preferred for numerical computation. Our new time-scaling strategy is hybrid in the sense that it is related to two coupled time-delay systems—one defined on the original time scale, in which the switching times are variable, the other defined on the new time scale, in which the switching times are fixed. This is different from the conventional time-scaling transformation widely used in the literature, which is not applicable to systems with time-delays. To demonstrate the effectiveness of the proposed approach, we solve four numerical examples. The results show that the costs obtained by our new approach are lower, when compared with those obtained by existing optimal control methods.

Changjun Yu
Central South University, Hunan, China, yuchangjun@126.com
Curtin University, Perth, Australia

Qun Lin
Curtin University, Perth, Australia, q.lin@curtin.edu.au

Ryan Loxton
Curtin University, Perth, Australia, r.loxton@curtin.edu.au

Kok Lay Teo, Corresponding author
Curtin University, Perth, Australia, k.l.teo@curtin.edu.au

Guoqiang Wang
Shanghai University of Engineering Science, Shanghai, China, guoq_wang@hotmail.com

1 Introduction

A dynamic system, whose evolution is influenced by the state and/or control
variables at some past time instants, is called a time-delay system. Such sys-
tems arise in many important applications in engineering and science, such as
medicine [1], aerospace engineering [2], and chemical reactor control [3–5].

The optimal control of time-delay systems has been an active area of re-
search for over three decades [6]. In particular, Many theoretical results are
now available in the literature, which include necessary optimality conditions
for time-delay optimal control problems [5,7]. However, except for some very
simple special cases, it is usually impossible to determine a closed-form an-
alytical expression for the optimal control. Thus, numerical methods are in-
dispensable for solving applied optimal control problems involving time-delay
systems [8–12].

The *Control Parameterization Technique* [12] is a popular numerical meth-
ods for solving optimal control problems, where the control function is approxi-
mated by a linear combination of basis functions, defined on a partition of fixed
subintervals. Applying this approximation scheme yields a finite-dimensional
approximation of the original optimal control problem, where the coefficients
in the approximating linear combination of basis functions are decision vari-
ables to be chosen optimally (see, for example, [13,14]). This approximate
problem is known as an optimal parameter selection problem. It is a nonlinear
optimization problem, which can be solved by using standard gradient-based
optimization techniques, such as sequential quadratic programming [15].

In the traditional control parameterization approach, the control is approx-
imated by a piecewise constant function, whose heights are decision variables
to be optimized. The switching times (the times at which the control switches
from one value to another) are typically equidistant, with no flexibility to
adaptively optimize their values. Thus, to obtain more accurate results, it is
usually necessary to choose a very fine partition of the time horizon. Conse-
quently, the finite-dimensional approximate optimization problem will consist
of a large number of decision variables. One way of reducing the number of
subintervals required is to also regard the control switching times as decision
variables, in addition to the control heights. However, it is known [16] that
gradient-based optimization methods are less desirable, when the switching
times are regarded as decision variables. The main reasons are: (i) the partial
derivatives of the cost and constraint functions with respect to the switching
times only exist, when the switching times are distinct; and (ii) numerical in-
tegration of the dynamic system over subintervals of variable length is harder

to implement numerically. For a detailed explanation on these two issues, see [16].

To overcome these difficulties, the time-scaling transformation [17,18] was developed in late 1990s to supplement the control parameterization technique. This transformation maps the variable switching times to fixed points in a new time horizon, thus yielding an equivalent optimization problem, in which the switching times are fixed. This new problem can be solved readily using standard gradient-based optimization techniques.

Although the time-scaling transformation has been successfully applied to a range of optimal control problems involving various system dynamics (e.g., switched systems [19], impulsive systems [20], discrete-valued control systems [21]), it has encountered difficulty for time-delay optimal control problems. This is because the time-scaling transformation affects the delay-length in a complex manner. In fact, under the time-scaling transformation, a fixed time-delay in the original time horizon becomes a variable time-delay in the new time horizon. Consequently, the resulting optimization problem can no longer be solved by using standard techniques. Hence, in a way, the time-scaling transformation complicates, rather than simplifies, the computation of optimal control problems involving time-delay systems.

In this paper, we propose a new time-scaling strategy to address this issue. We first apply the control parameterization technique to approximate the time-delay optimal control problem by a finite-dimensional optimization problem with canonical constraints, where the control heights and switching times are regarded as decision variables. Then, we develop a hybrid time-scaling transformation to map the variable switching times in the original time horizon to fixed points in a new time horizon. This hybrid time-scaling transformation only maps the current state and control vectors into the new time scale; the delayed state and control vectors remain in the original time scale. This means that the transformed dynamic system involves two different state and control functions: the state and control as functions of the original time variable, and the state and control as functions of the new time variable. Despite this complication, the new dynamic system has fixed switching times, and thus can be optimized using standard gradient-based optimization techniques.

In [8], the direct transcription is proposed to solve a class of time-delay optimal control problems, where the system dynamics and the cost function are discretized, giving rise to an approximate optimization problem. This optimization problem is solved sequentially on a sequence of successively refined grids. In [8], it is proposed that the nonlinear optimization problem be solved sequentially on a sequence of successively refined grids. However, the corresponding nonlinear optimization problem will eventually become a large scale nonlinear optimization problem, involving many equality constraints and decision variables, when the discretization becomes very fine. It is computationally expensive for solving such a large scale nonlinear optimization problem, especially with many nonlinear equality constraints. Nonetheless, this direct transcription approach has been applied successfully to solve time-delay optimal control problems.

The rest of the paper is organized as follows. In Section 2, we formulate the time-delay optimal control problem. In Section 3.1 and Section 3.2, we use the control parameterization method to yield a time-delay optimal parameter selection problem, which is a finite-dimensional approximation of the time-delay optimal control problem, where the switching times, as well as the control heights, are regarded as decision variables. Then, we introduce a new hybrid time-scaling transformation and use it to transform the time-delay optimal parameter selection problem with variable switching times into an equivalent problem, where the switching times are fixed. Section 3.3 is devoted to the derivation of the gradients of the cost function and the constraint functions of the equivalent problem. With these gradient formulas, the transformed time-delay optimal control problem can be solved by any gradient-based optimization technique. Section 4 contains numerical results, and finally, some concluding remarks are made in Section 5.

2 Problem Formulation

Consider the following time-delay system, defined on the fixed time interval $[-h, T]$:

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{f}(\boldsymbol{x}(t), \boldsymbol{x}(t-h), \boldsymbol{u}(t), \boldsymbol{u}(t-h)), \ t \in [0, T], \tag{1}$$

$$\boldsymbol{x}(t) = \boldsymbol{\phi}(t), \ t \in [-h, 0], \tag{2}$$

$$\boldsymbol{u}(t) = \boldsymbol{\varphi}(t), \ t \in [-h, 0[, \tag{3}$$

where $\boldsymbol{x} = [x_1, \ldots, x_n]^\top \in \mathbb{R}^n$ is the state vector, $\boldsymbol{u} = [u_1, \ldots, u_r]^\top \in \mathbb{R}^r$ is the control vector, h is a given time-delay satisfying $0 < h < T$, $\boldsymbol{f} : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^r \times \mathbb{R}^r \to \mathbb{R}^n$ and $\boldsymbol{\phi} : [-h, 0] \to \mathbb{R}^n$ are given continuously differentiable functions, and $\boldsymbol{\varphi} : [-h, 0[\to \mathbb{R}^r$ is a given piecewise continuous function. For notational simplicity, we only allow a single time-delay in (1). However, the results in this paper can be easily extended to the case of multiple time-delays, including the case, where the state and input delays are different.

Let U be a compact and convex subset of $\mathbb{R}^r$. A Borel measurable function $\boldsymbol{u} : [-h, T] \to \mathbb{R}^r$ is said to be an admissible control iff $\boldsymbol{u}(t) \in U$ for almost all $t \in [0, T]$, and $\boldsymbol{u}(t) = \boldsymbol{\varphi}(t)$ for all $t \in [-h, 0[$. Let $\mathscr{U}$ denote the class of all such admissible controls. For each $\boldsymbol{u} \in \mathscr{U}$, let $\boldsymbol{x}(\cdot|\boldsymbol{u})$ denote the corresponding absolutely continuous function satisfying the differential equation (1) almost everywhere on $]0, T]$, and the initial condition (2) everywhere on $[-h, 0]$. This function is called the solution of the dynamic system (1)-(3), corresponding to the control $\boldsymbol{u} \in \mathscr{U}$.

Our time-delay optimal control problem can be stated formally as follows: Given the dynamic system (1)-(3), find an admissible control $\boldsymbol{u} \in \mathscr{U}$ such that the cost functional

$$g_0(\boldsymbol{u}) = \Phi_0(\boldsymbol{x}(T|\boldsymbol{u})) + \int_0^T \mathcal{L}_0(\boldsymbol{x}(t|\boldsymbol{u}), \boldsymbol{x}(t-h|\boldsymbol{u}), \boldsymbol{u}(t)) dt \tag{4}$$

is minimized subject to the canonical equality constraints

$$g_i(\boldsymbol{u}) = \Phi_i(\boldsymbol{x}(T|\boldsymbol{u})) + \int_0^T \mathcal{L}_i(\boldsymbol{x}(t|\boldsymbol{u}), \boldsymbol{x}(t-h|\boldsymbol{u}), \boldsymbol{u}(t))dt = 0, \ i = 1, \ldots, N_e,$$
(5)

and the canonical inequality constraints

$$g_i(\boldsymbol{u}) = \Phi_i(\boldsymbol{x}(T|\boldsymbol{u})) + \int_0^T \mathcal{L}_i(\boldsymbol{x}(t|\boldsymbol{u}), \boldsymbol{x}(t-h|\boldsymbol{u}), \boldsymbol{u}(t))dt \geq 0,$$
$$i = N_e + 1, \ldots, N_e + N_i,$$
(6)

where $\Phi_i : \mathbb{R}^n \rightarrow \mathbb{R}$, $i = 0, \ldots, N_e + N_i$, and $\mathcal{L}_i : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^r \rightarrow \mathbb{R}$, $i = 0, \ldots, N_e + N_i$, are given real-valued functions. We refer to this problem as Problem (P).

Throughout this paper, we assume that the following conditions are satisfied.

A1. There exists a positive constant C such that

$$\|\boldsymbol{f}(\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \boldsymbol{\zeta}_1, \boldsymbol{\zeta}_2)\| \leq C(1 + \|\boldsymbol{\eta}_1\| + \|\boldsymbol{\eta}_2\| + \|\boldsymbol{\zeta}_1\| + \|\boldsymbol{\zeta}_2\|),$$
$$(\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \boldsymbol{\zeta}_1, \boldsymbol{\zeta}_2) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^r \times \mathbb{R}^r.$$

A2. $\mathcal{L}_i : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^r \rightarrow \mathbb{R}$, $i = 0, \ldots, N_e + N_i$, and $\Phi_i : \mathbb{R}^n \rightarrow \mathbb{R}$, $i = 0, \ldots, N_e + N_i$, are continuously differentiable with respect to each of their arguments.

3 Numerical Solution Procedure

3.1 Control Parameterization

We subdivide the time horizon $[0, T]$ into $q \geq 1$ subintervals. Let t_j, $j = 0, \ldots, q$, denote the end points of these subintervals, where $t_0 = 0$ and $t_q = T$. We impose the following constraints:

$$t_j - t_{j-1} \geq \epsilon, \ j = 1, \ldots, q,$$
(7)

where $\epsilon > 0$ is the minimum allowable subinterval duration. We now approximate the control in system (1)-(3) by a piecewise constant function, defined as follows:

$$\boldsymbol{u}(t) \approx \boldsymbol{u}^q(t) = \sum_{j=1}^q \boldsymbol{\delta}_j \chi_{[t_{j-1}, t_j[}(t), \ t \in [0, T],$$
(8)

where $\boldsymbol{\delta}_j \in U \subset \mathbb{R}^r$, $j = 1, \ldots, q$, and $\chi_{[a,b[}$ is the characteristic function, defined by

$$\chi_{[a,b[}(t) = \begin{cases} 1, & \text{if } t \in [a, b[, \\ 0, & \text{otherwise.} \end{cases}$$

Clearly, for each $t \in [0, T[$, either $t - h < 0$ or $t - h \in [t_{k-1}, t_k[$ for some integer $k = 1, \ldots, q$. If $t - h < 0$, then $\boldsymbol{u}(t - h) = \boldsymbol{\varphi}(t - h)$; if $t - h \in [t_{k-1}, t_k[$, then $\boldsymbol{u}(t - h) = \boldsymbol{\delta}_k$. Thus, for each $j = 1, \ldots, q$, substituting (8) into (1) yields the following new system, defined on the subinterval $[t_{j-1}, t_j[$:

$$
\dot{\boldsymbol{x}}(t) = \begin{cases} \boldsymbol{f}(\boldsymbol{x}(t), \boldsymbol{x}(t - h), \boldsymbol{\delta}_j, \boldsymbol{\delta}_k), & \text{if } t \in [t_{k-1} + h, t_k + h[, \\ & \quad \text{for some } k \in \{1, \ldots, q\} \\ \boldsymbol{f}(\boldsymbol{x}(t), \boldsymbol{\phi}(t - h), \boldsymbol{\delta}_j, \boldsymbol{\varphi}(t - h)), & \text{if } t < h. \end{cases} \quad (9)
$$

The approximate control (8) is completely defined by the switching times t_j, $j = 1, \ldots, q$, and the control parameters $\boldsymbol{\delta}_j$, $j = 1, \ldots, q$. Let Ξ denote the set of all vectors $\boldsymbol{\sigma} = [t_1, \ldots, t_{q-1}]^\top$ satisfying (7), and let Δ denote the set of all vectors $\boldsymbol{\delta} = [\boldsymbol{\delta}_1^\top, \ldots, \boldsymbol{\delta}_q^\top]^\top$ satisfying $\boldsymbol{\delta}_j \in U$, $j = 1, \ldots, q$. Furthermore, let $\boldsymbol{x}(\cdot | \boldsymbol{\delta}, \boldsymbol{\sigma})$ denote the solution of the system, defined by (2) and (9), corresponding to $(\boldsymbol{\delta}, \boldsymbol{\sigma}) \in \Delta \times \Xi$. Thus, after applying the control parameterization, the original optimal control problem becomes a time-delay optimal parameter selection problem, which is a finite-dimensional optimization problem. It is defined as follows: Given the initial conditions (2) and the dynamic system (9), find vectors $\boldsymbol{\delta} \in \Delta$ and $\boldsymbol{\sigma} \in \Xi$ such that the cost function

$$
\Phi_0(\boldsymbol{x}(T | \boldsymbol{\delta}, \boldsymbol{\sigma})) + \sum_{j=1}^q \int_{t_{j-1}}^{t_j} \mathcal{L}_0(\boldsymbol{x}(t | \boldsymbol{\delta}, \boldsymbol{\sigma}), \boldsymbol{x}(t - h | \boldsymbol{\delta}, \boldsymbol{\sigma}), \boldsymbol{\delta}_j) dt \quad (10)
$$

is minimized subject to the canonical equality constraints

$$
\Phi_i(\boldsymbol{x}(T | \boldsymbol{\delta}, \boldsymbol{\sigma})) + \sum_{j=1}^q \int_{t_{j-1}}^{t_j} \mathcal{L}_i(\boldsymbol{x}(t | \boldsymbol{\delta}, \boldsymbol{\sigma}), \boldsymbol{x}(t - h | \boldsymbol{\delta}, \boldsymbol{\sigma}), \boldsymbol{\delta}_j) dt = 0, \ i = 1, \ldots, N_e,
$$
$$(11)$$

and the canonical inequality constraints

$$
\Phi_i(\boldsymbol{x}(T | \boldsymbol{\delta}, \boldsymbol{\sigma})) + \sum_{j=1}^q \int_{t_{j-1}}^{t_j} \mathcal{L}_i(\boldsymbol{x}(t | \boldsymbol{\delta}, \boldsymbol{\sigma}), \boldsymbol{x}(t - h | \boldsymbol{\delta}, \boldsymbol{\sigma}), \boldsymbol{\delta}_j) dt \geq 0,
$$
$$i = N_e + 1, \ldots, N_e + N_i. \quad (12)$$

This problem is referred to as Problem $(P(q))$.

3.2 A Hybrid Time-scaling Transformation

To overcome the difficulties caused by variable switching times, the *time-scaling Transformation* was introduced in [18] to map the variable switching times to fixed points in a new time horizon. Thus, it gives rise to a new optimization problem, that is easier to solve. This transformation, however, is only applicable to systems without time-delays. Indeed, it turns out that for the class of time-delay optimal control problems considered in this paper, the time-scaling transformation causes the time-delay to become a function of the

subinterval durations in the original system. Because of this complication, it does not appear that the resulting optimization problem can be solved directly as such. We now introduce a hybrid time-scaling strategy to address this issue.

We define the time-scaling function, a one-to-one mapping from $[0, q]$ (the new time domain) to $[0, T]$ (the original time domain), as follows:

$$\mu(s|\boldsymbol{\theta}) = \begin{cases} T, & \text{if } s = q, \\ \sum_{j=1}^{\lfloor s \rfloor} \theta_j + \theta_{\lfloor s \rfloor + 1}(s - \lfloor s \rfloor), & \text{if } s \in [0, q[, \\ s, & \text{if } s < 0, \end{cases} \tag{13}$$

where $\theta_j = t_j - t_{j-1}$ is the length of the jth subinterval, and $\lfloor \cdot \rfloor$ denotes the floor function. Clearly, by (7),

$$\theta_j \geq \epsilon, \ j = 1, \ldots, q. \tag{14}$$

Moreover,

$$\theta_1 + \cdots + \theta_q = T. \tag{15}$$

Let Θ denote the set of all vectors $\boldsymbol{\theta} = [\theta_1, \ldots, \theta_q]^\top$ satisfying (14) and (15).

It is easy to see that

$$\frac{d\mu(s|\boldsymbol{\theta})}{ds} = \theta_{\lfloor s \rfloor + 1}, \ s \in [0, q] \backslash \{0, 1, \ldots, q\},$$

and

$$\mu(j|\boldsymbol{\theta}) = \theta_1 + \cdots + \theta_j = t_j, \ j = 1, \ldots, q.$$

This shows that the time-scaling transformation $t = \mu(s|\boldsymbol{\theta})$ maps $s = j$ to $t = t_j$. Let $\bar{\boldsymbol{x}}(s) = \boldsymbol{x}(\mu(s|\boldsymbol{\theta}))$. Then, applying the time-scaling transformation $t = \mu(s|\boldsymbol{\theta})$ gives the following new differential equation:

$$\dot{\bar{\boldsymbol{x}}}(s) = \frac{d}{ds}\big(\bar{\boldsymbol{x}}(s)\big) = \frac{d}{ds}\big(\boldsymbol{x}(\mu(s|\boldsymbol{\theta}))\big) = \theta_{\lfloor s \rfloor + 1} \frac{d\boldsymbol{x}(\mu(s|\boldsymbol{\theta}))}{dt}, \ s \in [0, q], \tag{16}$$

with the initial condition

$$\bar{\boldsymbol{x}}(s) = \boldsymbol{\phi}(s), \ s \in [-h, 0]. \tag{17}$$

Thus, by substituting (9) into (16) and (17), we obtain, for $\mu(s|\boldsymbol{\theta}) \in [t_{j-1}, t_j[$,

$$\dot{\bar{\boldsymbol{x}}}(s) = \begin{cases} \theta_{\lfloor s \rfloor + 1} \boldsymbol{f}(\bar{\boldsymbol{x}}(s), \boldsymbol{x}(\mu(s|\boldsymbol{\theta}) - h), \boldsymbol{\delta}_j, \boldsymbol{\delta}_k), & \text{if } \mu(s|\boldsymbol{\theta}) \in [t_{k-1} + h, t_k + h[, \\ \theta_{\lfloor s \rfloor + 1} \boldsymbol{f}(\bar{\boldsymbol{x}}(s), \boldsymbol{\phi}(\mu(s|\boldsymbol{\theta}) - h), \boldsymbol{\delta}_j, \boldsymbol{\varphi}(\mu(s|\boldsymbol{\theta}) - h)), & \text{if } \mu(s|\boldsymbol{\theta}) < h. \end{cases} \tag{18}$$

Note that the delayed argument $\mu(s|\boldsymbol{\theta}) - h$ in (18) depends on both s and $\boldsymbol{\theta}$. Let $s_{\text{delay}} = s_{\text{delay}}(s, \boldsymbol{\theta})$ and $t_{\text{delay}} = t_{\text{delay}}(s, \boldsymbol{\theta})$ denote the delayed time points in the new and original time horizons, respectively. Then, the relationship between s_{delay} and t_{delay} is given by the following fundamental equation:

$$\mu(s_{\text{delay}}|\boldsymbol{\theta}) = t_{\text{delay}} = \mu(s|\boldsymbol{\theta}) - h, \ s \in [0, q]. \tag{19}$$
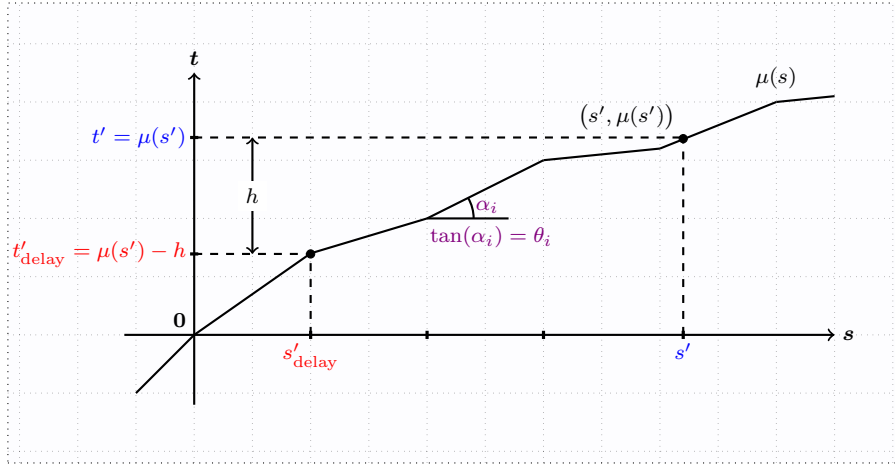
Fig. 1: The relationship between (s', t') and $(s'_{\text{delay}}, t'_{\text{delay}})$. Note that μ is a piecewise linear function, the slope of each linear segment $\tan(\alpha_i)$ corresponds to a duration variable θ_i.

This relationship is illustrated in Figure 1. Clearly, from Figure 1, whenever s', h, or $\boldsymbol{\theta}$ is changed, the corresponding time-delay, defined in the new time horizon s'_{delay}, will be changed as well.

Since $\mu(\cdot|\boldsymbol{\theta})$ is strictly increasing, $\mu(s|\boldsymbol{\theta}) \in [\mu(j-1|\boldsymbol{\theta}), \mu(j|\boldsymbol{\theta})[= [t_{j-1}, t_j[$ for all $s \in [j-1, j[$. Thus, using the new variable s_{delay}, (18) on the interval $[j-1, j[$ can be written as

$$
\dot{\bar{\boldsymbol{x}}}(s) = \begin{cases} \theta_j \boldsymbol{f}(\bar{\boldsymbol{x}}(s), \bar{\boldsymbol{x}}(s_{\text{delay}}), \boldsymbol{\delta}_j, \boldsymbol{\delta}_k), & \text{if } \mu(s_{\text{delay}}|\boldsymbol{\theta}) \in [t_{k-1}, t_k[, \\ \theta_j \boldsymbol{f}(\bar{\boldsymbol{x}}(s), \boldsymbol{\phi}(\mu(s_{\text{delay}}|\boldsymbol{\theta})), \boldsymbol{\delta}_j, \boldsymbol{\varphi}(\mu(s_{\text{delay}}|\boldsymbol{\theta}))), & \text{if } \mu(s_{\text{delay}}|\boldsymbol{\theta}) < 0. \end{cases}
$$
(20)

Equation (20) is a delay-differential equation, defined on the new time horizon $[0, q]$. Unlike (9), the switching points for (20) are fixed: they coincide with the integers $1, 2, \ldots, q-1$, and the new time-delay s_{delay} is now a variable in the new time horizon. The difficulty with solving (20) is that s_{delay} depends on s and the optimization parameters $\theta_j, j = 1, \ldots, q$. If $\mu(s_{\text{delay}}|\boldsymbol{\theta}) < 0$, it follows from (13) that $s_{\text{delay}} < 0$, and thus

$$
s_{\text{delay}} = \mu(s_{\text{delay}}|\boldsymbol{\theta}) = \mu(s|\boldsymbol{\theta}) - h. \tag{21}
$$

Hence, s_{delay} can be determined explicitly in this case. If $\mu(s_{\text{delay}}|\boldsymbol{\theta}) \geq 0$, on the other hand, it follows from (13) and (19) that s_{delay} satisfies the following equation:

$$
\sum_{j=1}^{\lfloor s \rfloor} \theta_j + \theta_{\lfloor s \rfloor + 1}(s - \lfloor s \rfloor) - h = \sum_{j=1}^{\lfloor s_{\text{delay}} \rfloor} \theta_j + \theta_{\lfloor s_{\text{delay}} \rfloor + 1}(s_{\text{delay}} - \lfloor s_{\text{delay}} \rfloor). \tag{22}
$$

To get value of the delayed state $\bar{\boldsymbol{x}}(s_{\text{delay}})$, we need to find the delay time s_{delay}. Note that s_{delay} is a variable, which is determined through solving (22) at each time point s. Thus, to calculate a single cost function value, it is required to solve this equation at least T/ϵ times, where T is the terminal time of the original time-delay system, and ϵ, which is usually taken as 10^{-3}, is the maximum step length, when solving the time-delay differential equations. Clearly, this process can be very time consuming, when a gradient-based optimization technique is used to solve the new problem, defined on the new time horizon. Thus, we need to find another way to obtain the value of $\bar{\boldsymbol{x}}(s_{\text{delay}})$ and then solve the new time-delay system (20).

To avoid the hassle of determining s_{delay} explicitly, we use the following identity:

$$\bar{\boldsymbol{x}}(s_{\text{delay}}) = \boldsymbol{x}(\mu(s_{\text{delay}}|\boldsymbol{\theta})) = \boldsymbol{x}(t_{\text{delay}}). \tag{23}$$

Substituting (19) and (23) into (20) yields the following dynamic system on the interval $[j-1, j[$:

$$\dot{\bar{\boldsymbol{x}}}(s) = \begin{cases} \theta_j \boldsymbol{f}(\bar{\boldsymbol{x}}(s), \boldsymbol{x}(t_{\text{delay}}), \boldsymbol{\delta}_j, \boldsymbol{\delta}_k), & \text{if } t_{\text{delay}} \in [t_{k-1}, t_k[, \\ \theta_j \boldsymbol{f}(\bar{\boldsymbol{x}}(s), \boldsymbol{\phi}(t_{\text{delay}}), \boldsymbol{\delta}_j, \boldsymbol{\varphi}(t_{\text{delay}})), & \text{if } t_{\text{delay}} < 0. \end{cases} \tag{24}$$

Note that t_{delay} (unlike s_{delay}) can be determined explicitly at all points, according to the equation $t_{\text{delay}} = \mu(s|\boldsymbol{\theta}) - h$. Thus, there is no difficulty in evaluating the delay in system (24). However, the right-hand side of (24) depends on two different state vectors: the state vector defined on the original time scale, and the state vector defined on the new time scale. Hence, our proposed time-scaling transformation is a hybrid transformation, that combines information from the new and old time scales.

Let $\bar{\boldsymbol{x}}(\cdot|\boldsymbol{\delta}, \boldsymbol{\theta})$ denote the solution of (17) and (24), corresponding to the pair $(\boldsymbol{\delta}, \boldsymbol{\theta}) \in \Delta \times \Theta$. For a given pair $(\boldsymbol{\delta}, \boldsymbol{\theta}) \in \Delta \times \Theta$, $\bar{\boldsymbol{x}}(\cdot|\boldsymbol{\delta}, \boldsymbol{\theta})$ can be determined as follows. For $s \leq 0$, $\bar{\boldsymbol{x}}(s|\boldsymbol{\delta}, \boldsymbol{\theta}) = \boldsymbol{\phi}(s)$, according to the initial condition (17). For $s > 0$, we need to determine $\bar{\boldsymbol{x}}(\cdot|\boldsymbol{\delta}, \boldsymbol{\theta})$ by solving the differential equation (24) numerically. When solving this differential equation, $\boldsymbol{x}(t_{\text{delay}})$ on the right-hand side is obtained using the previous values of $\bar{\boldsymbol{x}}(\cdot|\boldsymbol{\delta}, \boldsymbol{\theta})$. More specifically, to determine the state value at $t_{\text{delay}} = \mu(s|\boldsymbol{\theta}) - h$, we first find the unique past time points τ_i on the original time horizon such that $\tau_i \leq t_{\text{delay}} \leq \tau_{i+1}$. Such knot points must exist because $\mu(\cdot|\boldsymbol{\theta})$ is continuous and strictly increasing on $]-\infty, T]$. The value of $\boldsymbol{x}(t_{\text{delay}})$ can be obtained by interpolating the values of $\boldsymbol{x}(\tau_i)$ and $\boldsymbol{x}(\tau_{i+1})$. Figure 2 demonstrates the procedure to determine the value of $\boldsymbol{x}(t_{\text{delay}})$.

Remark 3.1 From a computational point of view, by using the traditional time-scaling transformation, solving a non-delayed problem requires storing an array of states $\{\bar{\boldsymbol{x}}(0), \ldots, \bar{\boldsymbol{x}}(q)\}$ and an array of times $\{0, \ldots, q\}$, and by applying the hybrid time-scaling transformation proposed in this paper, our solution for the delayed problem involves adding a second array of times $\{0, \ldots, T\}$. Thus, we try to find the correct value for the state by searching the array of transformed times (to find the non-delayed values of the state), and find the
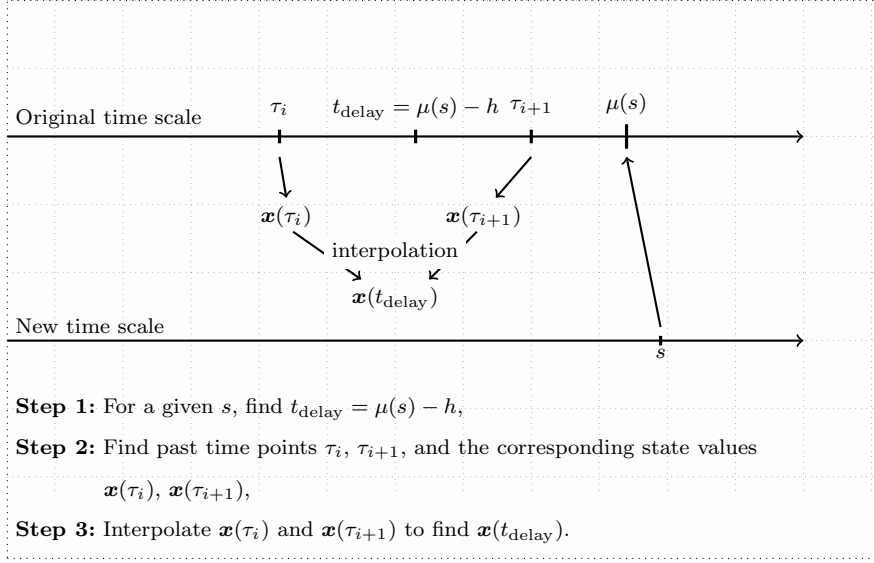
Step 1: For a given s, find $t_{\text{delay}} = \mu(s) - h$,

Step 2: Find past time points τ_i, τ_{i+1}, and the corresponding state values

   $\boldsymbol{x}(\tau_i)$, $\boldsymbol{x}(\tau_{i+1})$,

Step 3: Interpolate $\boldsymbol{x}(\tau_i)$ and $\boldsymbol{x}(\tau_{i+1})$ to find $\boldsymbol{x}(t_{\text{delay}})$.

Fig. 2: The procedure for determining $\boldsymbol{x}(t_{\text{delay}})$.

state by searching the array of original times (to find the delayed values of the state). In other words, we add a new "key" to the array of states, because performing the calculation in (22) can be time consuming.

After applying the hybrid time-scaling transformation, Problem $(P(q))$ becomes: Given the initial conditions (17) and the dynamic system (24), find vectors $\boldsymbol{\delta} \in \Delta$ and $\boldsymbol{\theta} \in \Theta$ such that the cost function

$$\bar{g}_0(\boldsymbol{\delta}, \boldsymbol{\theta}) = \Phi_0(\bar{\boldsymbol{x}}(q)) + \sum_{j=1}^{q} \int_{j-1}^{j} \theta_j \mathcal{L}_0(\bar{\boldsymbol{x}}(s), \boldsymbol{x}(t_{\text{delay}}), \boldsymbol{\delta}_j) ds \qquad (25)$$

is minimized subject to the canonical equality constraints

$$\bar{g}_i(\boldsymbol{\delta}, \boldsymbol{\theta}) = \Phi_i(\bar{\boldsymbol{x}}(q)) + \sum_{j=1}^{q} \int_{j-1}^{j} \theta_j \mathcal{L}_i(\bar{\boldsymbol{x}}(s), \boldsymbol{x}(t_{\text{delay}}), \boldsymbol{\delta}_j) ds = 0, \ i = 1, \ldots, N_e,$$
$$(26)$$

and the canonical inequality constraints

$$\bar{g}_i(\boldsymbol{\delta}, \boldsymbol{\theta}) = \Phi_i(\bar{\boldsymbol{x}}(q)) + \sum_{j=1}^{q} \int_{j-1}^{j} \theta_j \mathcal{L}_i(\bar{\boldsymbol{x}}(s), \boldsymbol{x}(t_{\text{delay}}), \boldsymbol{\delta}_j) ds \geq 0,$$
$$i = N_e + 1, \ldots, N_e + N_i, \qquad (27)$$

where $\bar{\boldsymbol{x}}(s) = \bar{\boldsymbol{x}}(s|\boldsymbol{\delta}, \boldsymbol{\theta})$ and $\boldsymbol{x}(t_{\text{delay}}) = \bar{\boldsymbol{x}}(s_{\text{delay}}|\boldsymbol{\delta}, \boldsymbol{\theta})$. We refer to this problem as Problem $(\bar{P}(q))$.

3.3 Gradient Computation

To solve Problem $(\bar{P}(q))$ using nonlinear optimization algorithms, we require the gradients of the cost and constraint functions. We first rewrite $\bar{g}_i$ in the following form:

$$\bar{g}_i(\boldsymbol{\delta}, \boldsymbol{\theta}) = \Phi_i(\bar{\boldsymbol{x}}(q)) + \int_0^q \frac{d\mu(s|\boldsymbol{\theta})}{ds}\hat{\mathcal{L}}_i(\bar{\boldsymbol{x}}(s), \boldsymbol{x}(t_{\text{delay}}), \boldsymbol{\delta})ds, \qquad (28)$$

where

$$\hat{\mathcal{L}}_i(\bar{\boldsymbol{x}}(s), \boldsymbol{x}(t_{\text{delay}}), \boldsymbol{\delta}) = \sum_{j=1}^q \mathcal{L}_i(\bar{\boldsymbol{x}}(s), \boldsymbol{x}(t_{\text{delay}}), \boldsymbol{\delta}_j)\chi_{[j-1,j)}(s).$$

The gradient of the state with respect to the duration vector $\boldsymbol{\theta}$ is given as a theorem stated below.

Theorem 3.1 *For each pair* $(\boldsymbol{\delta}, \boldsymbol{\theta}) \in \Delta \times \Theta,$

$$\frac{\partial \bar{\boldsymbol{x}}(s|\boldsymbol{\delta}, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \bar{\Lambda}(s|\boldsymbol{\delta}, \boldsymbol{\theta}), \ \ s \in [0, q].$$

Here, $\bar{\Lambda}(\cdot|\boldsymbol{\delta}, \boldsymbol{\theta})$ *is the solution of the following auxiliary dynamic system on each* $[j-1, j[$:

$$\begin{aligned}
\dot{\bar{\Lambda}}(s) &= \frac{\partial \hat{\boldsymbol{f}}^j(\bar{\boldsymbol{x}}(s), \boldsymbol{x}(t_{delay}), \boldsymbol{\theta}, \boldsymbol{\delta})}{\partial \bar{\boldsymbol{x}}}\bar{\Lambda}(s) \\
&+ \frac{\partial \hat{\boldsymbol{f}}^j(\bar{\boldsymbol{x}}(s), \boldsymbol{x}(t_{delay}), \boldsymbol{\theta}, \boldsymbol{\delta})}{\partial \boldsymbol{x}}\left[\bar{\Lambda}(s_{delay}) + \frac{\partial \boldsymbol{x}(t_{delay})}{\partial s_{delay}}\frac{\partial s_{delay}}{\partial \boldsymbol{\theta}}\right] \qquad (29) \\
&+ \frac{\partial \hat{\boldsymbol{f}}^j(\bar{\boldsymbol{x}}(s), \boldsymbol{x}(t_{delay}), \boldsymbol{\theta}, \boldsymbol{\delta})}{\partial \boldsymbol{\theta}}
\end{aligned}$$

with

$$\bar{\Lambda}(s) = \boldsymbol{0}, \ \ s \leq 0, \qquad (30)$$

where for $s \in [j-1, j[$,

$$\hat{\boldsymbol{f}}^j(\bar{\boldsymbol{x}}(s), \boldsymbol{x}(t_{delay}), \boldsymbol{\theta}, \boldsymbol{\delta}) = \begin{cases} \theta_j \boldsymbol{f}(\bar{\boldsymbol{x}}(s), \boldsymbol{x}(t_{delay}), \boldsymbol{\delta}_j, \boldsymbol{\delta}_k), & \text{if } t_{delay} \in [t_{k-1}, t_k[, \\ \theta_j \boldsymbol{f}(\bar{\boldsymbol{x}}(s), \boldsymbol{x}(t_{delay}), \boldsymbol{\delta}_j, \boldsymbol{\varphi}(t_{delay})), & \text{if } t_{delay} < 0. \end{cases}$$

Proof The proof of the theorem is given in the Appendix. □

Theorem 3.2 *The gradient of* $\bar{g}_i(\boldsymbol{\delta}, \boldsymbol{\theta})$ *for each* $i = 0, 1, \ldots, N_e + N_i$ *with respect to* $\boldsymbol{\theta}$ *is given by*

$$\begin{aligned}
\frac{\partial \bar{g}_i(\boldsymbol{\delta}, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} &= \frac{\partial \Phi_i(\bar{\boldsymbol{x}}(q))}{\partial \bar{\boldsymbol{x}}} \cdot \frac{\partial \bar{\boldsymbol{x}}(q)}{\partial \boldsymbol{\theta}} + \int_0^q \left\{ \left[\frac{\partial \hat{\mathcal{L}}_i(\bar{\boldsymbol{x}}(s), \boldsymbol{x}(t_{delay}), \boldsymbol{\delta})}{\partial \bar{\boldsymbol{x}}} \cdot \frac{\partial \bar{\boldsymbol{x}}(s)}{\partial \boldsymbol{\theta}} \right. \right. \\
&+ \left. \frac{\partial \hat{\mathcal{L}}_i(\bar{\boldsymbol{x}}(s), \boldsymbol{x}(t_{delay}), \boldsymbol{\delta})}{\partial \boldsymbol{x}} \cdot \frac{\partial \boldsymbol{x}(t_{delay})}{\partial \boldsymbol{\theta}} \right] \frac{d\mu(s|\boldsymbol{\theta})}{ds} \qquad (31) \\
&+ \left. \hat{\mathcal{L}}_i(\bar{\boldsymbol{x}}(s), \boldsymbol{x}(t_{delay}), \boldsymbol{\delta}) \cdot \frac{\partial}{\partial \boldsymbol{\theta}}\left(\frac{d\mu(s|\boldsymbol{\theta})}{ds} \right) \right\} ds.
\end{aligned}$$

Proof The proof follows from applying the chain rule to (28) and Theorem 3.1. □

Remark 3.2

(i) To numerically solve the auxiliary dynamic system (29)-(30) without using the explicit form of s_{delay}, we let $\Lambda(\cdot)$ be the corresponding auxiliary state, defined on the original time scale. Then, it follows from (21) that $\Lambda(t_{\text{delay}}) = \bar{\Lambda}(s_{\text{delay}})$, and hence $\bar{\Lambda}(s_{\text{delay}})$ can be obtained by applying a similar interpolation procedure, as suggested in Section 3.2 for the computation of $\bar{\boldsymbol{x}}(s_{\text{delay}})$.

(ii) For the computation of $\partial \boldsymbol{x}(t_{\text{delay}})/\partial s_{\text{delay}}$, we note from (23) that it equals to $\dot{\bar{\boldsymbol{x}}}(s_{\text{delay}})$ and thus can be replaced by $\dot{\boldsymbol{x}}(t_{\text{delay}})$, which can be obtained by applying a similar procedure, as we depicted in Figure 2 for the computation of $\boldsymbol{x}(t_{\text{delay}})$.

(iii) For the computation of $\partial s_{\text{delay}}/\partial \boldsymbol{\theta}$, taking the partial derivative of both sides of (19) with respect to $\boldsymbol{\theta}$ gives

$$\frac{\partial \mu(s_{\text{delay}})}{\partial \boldsymbol{\theta}} + \frac{\partial \mu(s_{\text{delay}})}{\partial s} \frac{\partial s_{\text{delay}}}{\partial \boldsymbol{\theta}} = \frac{\partial \mu(s)}{\partial \boldsymbol{\theta}}. \tag{32}$$

Then, by rearranging (32), we have

$$\frac{\partial s_{\text{delay}}}{\partial \boldsymbol{\theta}} = \left\{ \frac{\partial \mu(s)}{\partial \boldsymbol{\theta}} - \frac{\partial \mu(s_{\text{delay}})}{\partial \boldsymbol{\theta}} \right\} \div \frac{\partial \mu(s_{\text{delay}})}{\partial s}. \tag{33}$$

(iv) The value of $\partial \mu(s_{\text{delay}})/\partial \boldsymbol{\theta}$ and $\partial \mu(s_{\text{delay}})/\partial s$ can also be obtained by interpolation without using the explicit form of s_{delay}.

The gradient of the state with respect to $\boldsymbol{\delta}$ is given within next theorem.

Theorem 3.3 *For each pair* $(\boldsymbol{\delta}, \boldsymbol{\theta}) \in \Delta \times \Theta$,

$$\frac{\partial \bar{\boldsymbol{x}}(s|\boldsymbol{\delta}, \boldsymbol{\theta})}{\partial \boldsymbol{\delta}} = \bar{\Upsilon}(s|\boldsymbol{\delta}, \boldsymbol{\theta}), \ \ s \in [0, q],$$

where $\bar{\Upsilon}(\cdot|\boldsymbol{\delta}, \boldsymbol{\theta})$ *is the solution of the following auxiliary dynamic system on each interval* $[j - 1, j[$:

$$\dot{\bar{\Upsilon}}(s) = \frac{\partial \hat{\boldsymbol{f}}^j\big(\bar{\boldsymbol{x}}(s), \boldsymbol{x}(t_{delay}), \boldsymbol{\theta}, \boldsymbol{\delta}\big)}{\partial \bar{\boldsymbol{x}}} \bar{\Upsilon}(s) + \frac{\partial \hat{\boldsymbol{f}}^j\big(\bar{\boldsymbol{x}}(s), \boldsymbol{x}(t_{delay}), \boldsymbol{\theta}, \boldsymbol{\delta}\big)}{\partial \boldsymbol{x}} \bar{\Upsilon}(s_{delay})$$
$$+ \frac{\partial \hat{\boldsymbol{f}}^j\big(\bar{\boldsymbol{x}}(s), \boldsymbol{x}(t_{delay}), \boldsymbol{\theta}, \boldsymbol{\delta}\big)}{\partial \boldsymbol{\delta}}, \tag{34}$$

with the initial condition

$$\bar{\Upsilon}(s) = \boldsymbol{0}, \ \ s \leq 0. \tag{35}$$

Proof The proof is similar to the proof of Theorem 3.1, and hence is omitted. □

Theorem 3.4 *For each* $i = 0, 1, \ldots, N_e + N_i$, *the gradient of* $\bar{g}_i(\boldsymbol{\delta}, \boldsymbol{\theta})$ *with respect to* $\boldsymbol{\delta}$ *is given by*

$$\frac{\partial \bar{g}_i(\boldsymbol{\delta}, \boldsymbol{\theta})}{\partial \boldsymbol{\delta}} = \frac{\partial \Phi_i(\bar{\boldsymbol{x}}(q))}{\partial \bar{\boldsymbol{x}}} \cdot \frac{\partial \bar{\boldsymbol{x}}(q)}{\partial \boldsymbol{\delta}} + \int_0^q \left\{ \left[\frac{\partial \hat{\mathcal{L}}_i(\bar{\boldsymbol{x}}(s), \boldsymbol{x}(t_{delay}), \boldsymbol{\delta})}{\partial \bar{\boldsymbol{x}}} \cdot \frac{\partial \bar{\boldsymbol{x}}(s)}{\partial \boldsymbol{\delta}} \right. \right.$$
$$\left. \left. + \frac{\partial \hat{\mathcal{L}}_i(\bar{\boldsymbol{x}}(s), \boldsymbol{x}(t_{delay}), \boldsymbol{\delta})}{\partial \boldsymbol{x}} \cdot \frac{\partial \boldsymbol{x}(t_{delay})}{\partial \boldsymbol{\delta}} \right] \frac{d\mu(s|\boldsymbol{\theta})}{ds} \right\} ds.$$

Proof The proof follows from applying the chain rule to (28) and Theorem 3.3. □

Recall that Problem $(\bar{P}(q))$ is an optimal parameter selection problem. Theorem 3.2 and Theorem 3.4 give the gradients of the cost and constraint functions in Problem $(\bar{P}(q))$ with respect to $\boldsymbol{\theta}$ and $\boldsymbol{\delta}$, respectively. On this basis, we can use any of the existing nonlinear programming software packages, which is gradient-based, to solve Problem $(\bar{P}(q))$. Examples include FMINCON in MATLAB or NLPQLP in FORTRAN. In next section, we will demonstrate the effectiveness of this approach through solving four numerical examples.

4 Numerical Results

4.1 Problem 1: Optimal Control Problem with both State and Control Delay

Consider the following time-delay optimal control problem, which includes time-delays in both the state and the control variables:

$$\min g_0(\boldsymbol{u}) = \frac{1}{2} \boldsymbol{x}(2)^\top S \boldsymbol{x}(2) + \frac{1}{2} \int_0^2 \left[\boldsymbol{x}(t)^\top Q \boldsymbol{x}(t) + \boldsymbol{u}(t)^\top R \boldsymbol{u}(t) \right] dt, \quad (36)$$

where

$$S = \begin{bmatrix} 1 & 2 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix}, \quad R = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad Q = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix},$$

subject to the time-delay dynamic system

$$\dot{x}_1(t) = -2x_1(t)^2 + x_1(t)x_2(t - 0.1) + 2x_2(t) - u_1(t)u_2(t - 0.1),$$
$$\dot{x}_2(t) = -2x_2(t)^2 - x_1(t - 0.1) + 2x_3(t) + u_2(t),$$
$$\dot{x}_3(t) = -x_3(t)^3 - x_1(t)x_2(t) - x_2(t - 0.1)u_2(t) + u_1(t - 0.1) + 2u_3(t),$$
$$\dot{x}_4(t) = -x_4(t)^2 + x_2(t)x_3(t) - 2x_3(t - 0.1) + 2u_4(t),$$

the initial conditions

$$x_i(t) = 1, \quad t \in [-0.1, 0], \quad i = 1, \ldots, 4,$$

Table 1: Optimal cost and constraint values for Problem 1 (variable subinterval lengths).

Number of Subintervals	$g_0(\boldsymbol{u}^{q,\star})$	Constraint values at $\boldsymbol{u}^{q,\star}$
$q = 8$	1.10082	$g_1(\boldsymbol{u}^{q,\star}) = 3.9596$ $g_2(\boldsymbol{u}^{q,\star}) = 0.0312$
$q = 6$	1.10484	$g_1(\boldsymbol{u}^{q,\star}) = 3.9072$ $g_2(\boldsymbol{u}^{q,\star}) = 0.0373$
$q = 4$	1.11843	$g_1(\boldsymbol{u}^{q,\star}) = 3.9214$ $g_2(\boldsymbol{u}^{q,\star}) = 0.0392$

Table 2: Optimal cost and constraint values for Problem 1 (equally-spaced subintervals).

Number of Subintervals	$g_0(\boldsymbol{u}^{q,\star})$	Constraint values at $\boldsymbol{u}^{q,\star}$
$q = 8$	1.10518	$g_1(\boldsymbol{u}^{q,\star}) = 3.9790$ $g_2(\boldsymbol{u}^{q,\star}) = 0.0077$
$q = 6$	1.11337	$g_1(\boldsymbol{u}^{q,\star}) = 3.9427$ $g_2(\boldsymbol{u}^{q,\star}) = 0.0092$
$q = 4$	1.13708	$g_1(\boldsymbol{u}^{q,\star}) = 3.9329$ $g_2(\boldsymbol{u}^{q,\star}) = 0.0043$

$$u_i(t) = 1, \quad t \in [-0.1, 0[, \quad i = 1, \ldots, 4,$$

the terminal inequality constraints

$$g_1(\boldsymbol{u}) = 4.0 - x_1(2)^2 - x_2(2)^2 - x_3(2)^2 - x_4(2)^2 \geq 0,$$
$$g_2(\boldsymbol{u}) = x_1(2)^2 + x_2(2)^2 + x_3(2)^2 + x_4(2)^2 - 0.002 \geq 0,$$

and the control constraints

$$-3 \leq u_i(t) \leq 1, \; t \in [0, 2], \; i = 1, \ldots, 4.$$

In [12], this problem was solved by applying the traditional control parameterization technique, with the planning horizon $[0, 2]$ being devided into $q = 20$ equally-spaced subintervals. The optimal control obtained yields an optimal cost of $g_0(\boldsymbol{u}^{\star}) = 1.10297$, and the corresponding constraint values are

$$g_1(\boldsymbol{u}^{\star}) = 3.9605, \quad g_2(\boldsymbol{u}^{\star}) = 0.0019.$$

To compare, we choose $q = 4, 6$, and 8, and then apply our new approach. The results from our new method and the traditional control parameterization method are listed in $Tables$ 1 and 2.

Clearly, from $Tables$ 1 and 2, when using the same number of subintervals, our new method always results in a better cost value, when compared with the conventional control parameterization technique. Again, this is expected, as the new approach provides the added flexibity to optimize the control switching times.

4.2 Problem 2: Continuous Stirred Tank Reactor

Consider the following time-delay optimal control problem, which comes from a continuous stirred tank reactor system [8], where the time-delays in the state and control are different:

$$\min g_0(\boldsymbol{u}) = \int_0^{0.2} \|\boldsymbol{x}(t)\|_2^2 + 0.01\|\boldsymbol{u}(t)\|_2^2 dt, \qquad (37)$$

subject to the time-delay dynamic system

$$\dot{x}_1(t) = -x_1(t) - R(t),$$
$$\dot{x}_2(t) = -x_2(t) + 0.9u_2(t - 0.02) + 0.1u_2(t),$$
$$\dot{x}_3(t) = -2x_3(t) + 0.25R(t) - 1.05u_1(t)x_3(t - 0.015),$$

where

$$R(t) = (1 + x_1(t))(1 + x_2(t)) \exp\left[\frac{25x_3(t)}{1 + x_3(t)}\right],$$

and the initial state and control are given by:

$$x_3(t) = -0.02, \quad -r \le t \le 0,$$
$$u_2(t) = 1, \quad -s \le t \le 0,$$
$$\boldsymbol{x}(0) = [0.49, -0.0002, -0.02]^\top.$$

Since the terminal time for this system is relatively small, we choose the number of intervals to be $q = 1$. In this case, the control function without using the proposed hybrid time-scaling transformation will be the same as the one using hybrid time-scaling transformation. The optimal control obtained is

$$\boldsymbol{u}^\star(t) = [0.0624, 0.0098]^\top, \ t \in]0, 2],$$

with a cost of $g_0(\boldsymbol{u}^\star) = 0.0213$ with a truncation error of 7.29E-10. Whereas in [8], the cost of 0.0213 is obtained with an error of 1.42E-8. Figure 3 gives the optimal states trajectory, corresponding to the optimal control.

4.3 Problem 3: LQR Optimal Control with State Delay

Consider the following time-delay optimal control problem [11]:

$$\min g_0(\boldsymbol{u}) = \frac{1}{2}\boldsymbol{x}^\top(t_f)S\boldsymbol{x}(t_f) + \frac{1}{2}\int_0^{t_f} \boldsymbol{x}^\top Q\boldsymbol{x} + \boldsymbol{u}^\top R\boldsymbol{u} dt, \qquad (38)$$

subject to the time-delay dynamic system

$$\dot{x}(t) = A_1(t)\boldsymbol{x}(t) + A_2(t)\boldsymbol{x}(t - h) + B(t)\boldsymbol{u}(t), \qquad (39)$$
$$\boldsymbol{x}(t) = [1, 0]^\top, \ h \le t \le 0, \qquad (40)$$
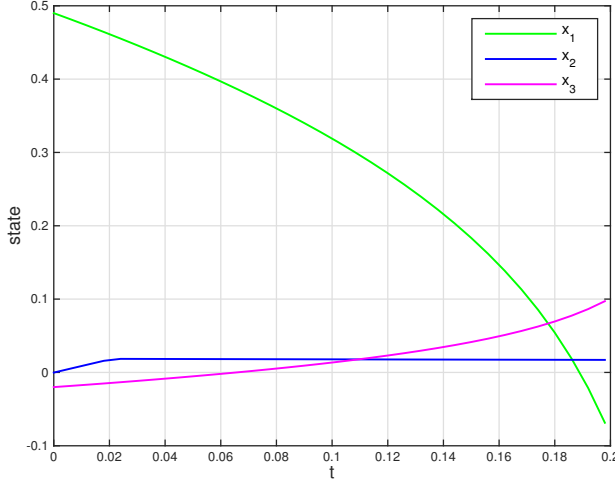
Fig. 3: Optimal state for Problem 2.

Table 3: Parameters in Problem 3

a	b	c	h	t_f	Q	R	S
0.2	0.5	0.2	1	2	$I_{2\times2}$	$I_{2\times2}$	$10^4 I_{2\times2}$

where

$$A_1(t) = \begin{bmatrix} 0 & 1 \\ -4\pi^2(a + c\cos 2\pi t) & 0 \end{bmatrix}, \ A_2(t) = \begin{bmatrix} 0 & 0 \\ -4\pi^2 b\cos 2\pi t & 0 \end{bmatrix}, \ B(t) = \begin{bmatrix} 0 \\ 1 \end{bmatrix},$$

and the parameters of the problem are in Table 3:

By applying the proposed hybrid time-scaling transformation to the time-delay system (39), and let $q = 5$, we obtain a cost of $g_0(\boldsymbol{u}^\star) = 3.4183$. However, if the time horizon is equally spaced, the corresponding cost value is 7.0065. The two optimal control functions are given in Figure 4, and the corresponding optimal control trajectories are given in Figure 5 and Figure 6. The results again show that the proposed method can achieve a much better cost, when compared with the traditional control parameterization technique.

It is noted that in [11], the optimal cost value is not provided. However, by comparing the control function, we have obtained a much lower cost.

4.4 Problem 4: A Variation of Problem 3

Consider a version of Problem 3 as considered in [8], where the cost function and the dynamic system are the same as those in Problem 3, except that

$$A_2(t) = \begin{bmatrix} 0 & 1 \\ 4\pi^2 b\cos 2\pi t & 0 \end{bmatrix}$$
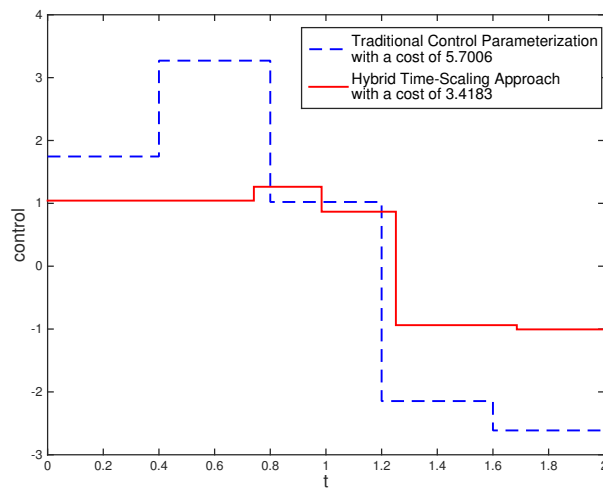
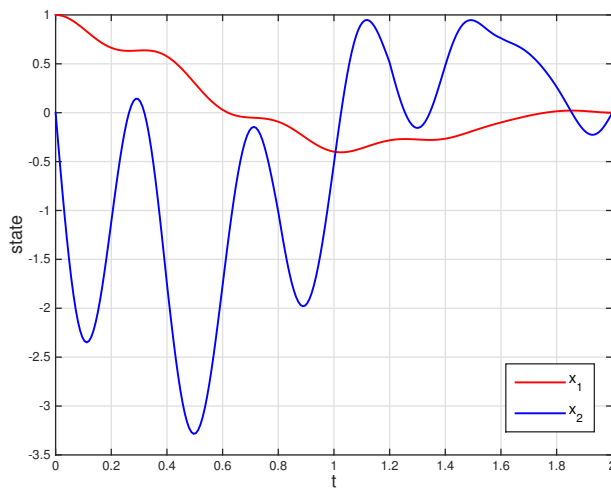Fig. 4: Optimal control for Problem 3 — $q = 5$.

Fig. 5: Optimal state trajectory for Problem 3 — using Traditional Control Parameterization with $q = 5$.

and the terminal time is $t_f = 50$. Note that the dynamic system is very unstable, the state vector will go to infinity if there is no control applied to the system. By choosing $q = 21$, we obtain a cost of $g_0(\boldsymbol{u}^\star) = 4.3554$. However, the traditional control parameterization technique fails to solve this problem.
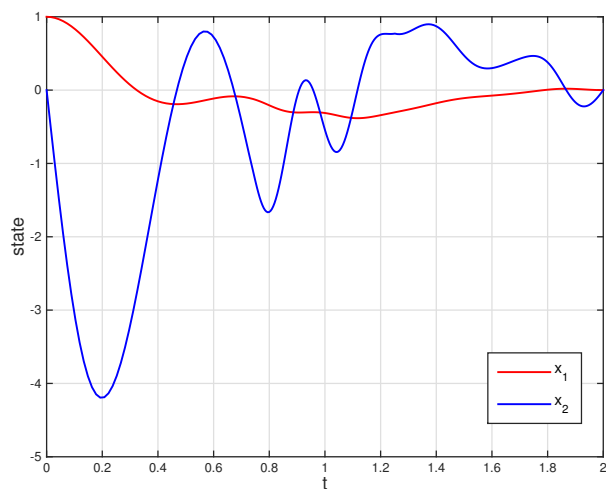
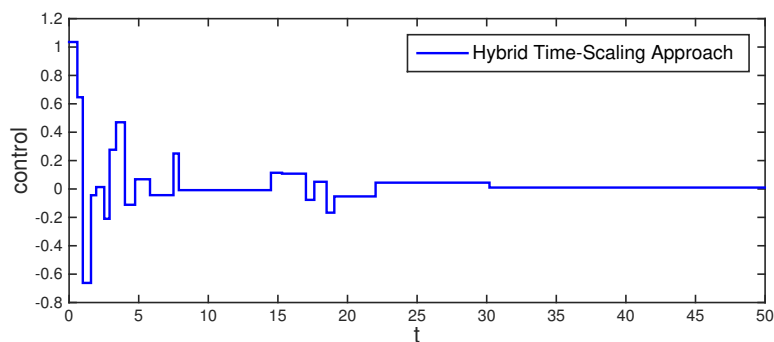Fig. 6: Optimal state trajectory for Problem 3 — using Hybrid Time-Scaling transformation with $q = 5$.

Fig. 7: Optimal control for Problem 4 — $q = 21$.

The obtained optimal control and the corresponding state are given in Figure 7 and Figure 8.

In [8], the optimal cost value is not provided. However, by comparing the control functions, clearly, we have obtained a lower cost.

5 Conclusions

In this paper, we develop a hybrid time-scaling transformation for time-delay optimal control problems. On this basis, a new computational approach is proposed to solve the time-delay optimal control problems. Although we only
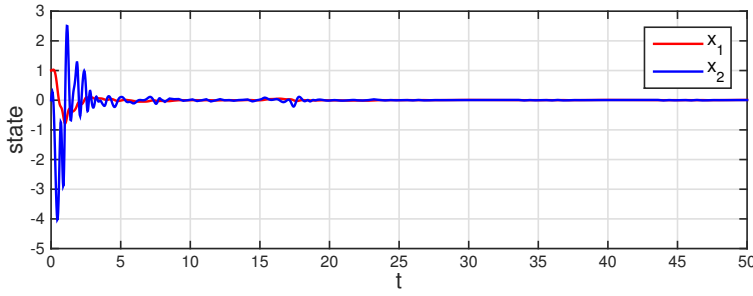
Fig. 8: Optimal state trajectory for Problem 4 — using Hybrid Time-Scaling transformation with $q = 21$.

demonstrate the approach for dynamic systems with a single state delay and a single control delay, the approach can be extended to systems with multiple state and input delays. Numerical results show that the proposed approach is highly effective, when compared with those obtained by using other existing method.

Appendix

Proof of Theorem 3.1.

Proof Let $\boldsymbol{\delta}$ and $r \in \{1, \ldots, q\}$ be arbitrary but fixed, and let $\boldsymbol{e}^r$ be the rth unit vector in $\mathbb{R}^q$. Then,

$$\frac{\partial \bar{\boldsymbol{x}}(s)}{\partial \theta_r} = \lim_{\xi \to 0} \frac{\bar{\boldsymbol{x}}(s|\boldsymbol{\delta}, \boldsymbol{\theta}^\xi) - \bar{\boldsymbol{x}}(s|\boldsymbol{\delta}, \boldsymbol{\theta})}{\xi}, \tag{41}$$

where $\boldsymbol{\theta}^\xi = \boldsymbol{\theta} + \xi \boldsymbol{e}^r$, s_{delay}^ξ and t_{delay}^ξ are the corresponding delayed time points in the new time horizon and the original time horizon such that

$$\mu(s_{\text{delay}}^\xi | \boldsymbol{\theta}^\xi) = t_{\text{delay}}^\xi = \mu(s|\boldsymbol{\theta}^\xi) - h, \ s \in [0, q].$$

Now, we will prove the theorem in the following steps.

Step 1: Preliminaries

For each real number $\xi \in \mathbb{R}$, let $\bar{\boldsymbol{x}}^\xi$ denote the function $\bar{\boldsymbol{x}}(\cdot | \boldsymbol{\delta}, \boldsymbol{\theta}^\xi)$. Then, it follows from (20) that, for each $\xi \in \mathbb{R}$,

$$\bar{\boldsymbol{x}}^\xi(s) = \bar{\boldsymbol{x}}(s) + \int_0^s \boldsymbol{F}^\xi(t) dt, \quad s \in [0, q], \tag{42}$$

where $\bar{\boldsymbol{x}}(s)$ denotes the function $\bar{\boldsymbol{x}}\big(\,\cdot\,|\boldsymbol{\delta},\boldsymbol{\theta}\big)$, and $\boldsymbol{F}^{\xi}$ is defined as follows:

$$
\boldsymbol{F}^{\xi}(s) := \begin{cases} \theta^{\xi}_{\lfloor s \rfloor+1}\boldsymbol{f}\big(\bar{\boldsymbol{x}}(s|\boldsymbol{\delta},\boldsymbol{\theta}^{\xi}),\bar{\boldsymbol{x}}(s^{\xi}_{\mathrm{delay}}|\boldsymbol{\delta},\boldsymbol{\theta}^{\xi}),\boldsymbol{\delta}\big), & \text{if } s_{\mathrm{delay}} \geq 0, \\[2mm] \theta^{\xi}_{\lfloor s \rfloor+1}\boldsymbol{f}\big(\bar{\boldsymbol{x}}(s|\boldsymbol{\delta},\boldsymbol{\theta}^{\xi}),\bar{\boldsymbol{x}}(s^{\xi}_{\mathrm{delay}}|\boldsymbol{\delta},\boldsymbol{\theta}^{\xi}),\boldsymbol{\delta},\varphi(t^{\xi}_{\mathrm{delay}})\big), & \text{if } s_{\mathrm{delay}} < 0. \end{cases}
$$

Define

$$
\begin{aligned}
\Gamma^{\xi}(s) &= \bar{\boldsymbol{x}}(s|\boldsymbol{\delta},\boldsymbol{\theta}^{\xi}) - \bar{\boldsymbol{x}}(s|\boldsymbol{\delta},\boldsymbol{\theta}) \\
&= \int_0^s \big(\boldsymbol{F}^{\xi} - \boldsymbol{F}^0\big)dt.
\end{aligned} \tag{43}
$$

Applying the mean value theorem, we have, for $s \in [0,q]$,

$$
\begin{aligned}
&\boldsymbol{F}^{\xi}(s) - \boldsymbol{F}^0(s) \\
&= \int_0^1 \Bigg\{ \frac{\partial \boldsymbol{f}\big(\bar{\boldsymbol{x}}+\eta\Gamma^{\xi}(s),\bar{\boldsymbol{x}}_d+\eta(\bar{\boldsymbol{x}}^{\xi}_d-\bar{\boldsymbol{x}}_d),\boldsymbol{\theta}+\eta\xi\boldsymbol{e}^r,\boldsymbol{\delta}\big)}{\partial \bar{\boldsymbol{x}}}\Gamma^{\xi}(s) \\
&\quad + \frac{\partial \boldsymbol{f}\big(\bar{\boldsymbol{x}}+\eta\Gamma^{\xi}(s),\bar{\boldsymbol{x}}_d+\eta(\bar{\boldsymbol{x}}^{\xi}_d-\bar{\boldsymbol{x}}_d),\boldsymbol{\theta}+\eta\xi\boldsymbol{e}^r,\boldsymbol{\delta}\big)}{\partial \bar{\boldsymbol{x}}_d}(\bar{\boldsymbol{x}}^{\xi}_d-\bar{\boldsymbol{x}}_d) \\
&\quad + \frac{\partial \boldsymbol{f}\big(\bar{\boldsymbol{x}}+\eta\Gamma^{\xi}(s),\bar{\boldsymbol{x}}_d+\eta(\bar{\boldsymbol{x}}^{\xi}_d-\bar{\boldsymbol{x}}_d),\boldsymbol{\theta}+\eta\xi\boldsymbol{e}^r,\boldsymbol{\delta}\big)}{\partial \theta_r}\xi \Bigg\}d\eta,
\end{aligned} \tag{44}
$$

and

$$
\begin{aligned}
\bar{\boldsymbol{x}}^{\xi}_d - \bar{\boldsymbol{x}}^0_d &= \bar{\boldsymbol{x}}(s^{\xi}_{\mathrm{delay}}|\boldsymbol{\delta},\boldsymbol{\theta}^{\xi}) - \bar{\boldsymbol{x}}(s_{\mathrm{delay}}|\boldsymbol{\delta},\boldsymbol{\theta}) \\
&= \bar{\boldsymbol{x}}(s^{\xi}_{\mathrm{delay}}|\boldsymbol{\delta},\boldsymbol{\theta}^{\xi}) - \bar{\boldsymbol{x}}(s_{\mathrm{delay}}|\boldsymbol{\delta},\boldsymbol{\theta}^{\xi}) + \bar{\boldsymbol{x}}(s_{\mathrm{delay}}|\boldsymbol{\delta},\boldsymbol{\theta}^{\xi}) - \bar{\boldsymbol{x}}(s_{\mathrm{delay}}|\boldsymbol{\delta},\boldsymbol{\theta}) \\
&= \bar{\boldsymbol{x}}(s^{\xi}_{\mathrm{delay}}|\boldsymbol{\delta},\boldsymbol{\theta}^{\xi}) - \bar{\boldsymbol{x}}(s_{\mathrm{delay}}|\boldsymbol{\delta},\boldsymbol{\theta}^{\xi}) + \Gamma^{\xi}(s_{\mathrm{delay}})
\end{aligned}
$$

From Assumption **A1**, it follows that the state set $\{\bar{\boldsymbol{x}}^{\xi}(s) : \xi \in [-a,a]\}$ is equibounded on $[0,q]$, where $a > 0$ is a fixed small real number. Hence there exists a real number $C_1 > 0$ such that for each $\xi \in [-a,a]$,

$$
\bar{\boldsymbol{x}}^{\xi}(s) \in \mathcal{N}_n(C_1), \ \ s \in [0,q],
$$

where $\mathcal{N}_n(C_1)$ denotes the closed ball in $\mathbb{R}^n$ of radius C_1 centered at the origin. Note that, $\mathcal{N}_n(C_1)$ is convex, thus, for each $\xi \in [-a,a]$,

$$
\bar{\boldsymbol{x}}(s) + \eta\Gamma^{\xi}(s) \in \mathcal{N}_n(C_1), \quad s \in [0,q], \ \eta \in [0,1].
$$

Furthermore, it is easy to see that for each $\xi \in [-a,a]$,

$$
\boldsymbol{\theta} + \eta\xi\boldsymbol{e}^r \in \mathcal{N}_q(C_2), \quad \eta \in [0,1],
$$

where $C_2 := |\boldsymbol{\theta}|_q + a$. Recall from Assumption **A2** that $\partial \boldsymbol{f}/\partial \boldsymbol{x}$ and $\partial \boldsymbol{f}/\partial \theta_r$ are continuous. Hence, it follows from the compactness of $[0,T]$, $\mathcal{V}$, $\mathcal{N}_n(C_1)$ and $\mathcal{N}_q(C_2)$ and the definitions of $z(s|\boldsymbol{\theta})$ and ϕ that there exists a real number $C_3 > 0$ such that, for each $\xi \in [-a,a]$,

$$
\left|\frac{\partial \boldsymbol{f}^{\xi}_{\eta}}{\partial \bar{\boldsymbol{x}}}\right|_{n \times n} \leq C_3, \quad s \in [0,q], \ \eta \in [0,1],
$$

$$\left|\frac{\partial \boldsymbol{f}_\eta^\xi}{\partial \bar{\boldsymbol{x}}_d}\right|_{n\times n} \leq C_3, \quad s \in [0, q], \; \eta \in [0, 1],$$

$$\left|\frac{\partial \boldsymbol{f}_\eta^\xi}{\partial \theta_r}\right|_n \leq C_3, \quad s \in [0, q], \; \eta \in [0, 1],$$

$$\left|\frac{\partial \phi_\eta^\xi}{\partial t}\right|_n \leq C_3, \quad s \in [0, q], \; \eta \in [0, 1],$$

where $\boldsymbol{f}_\eta^\xi$ denotes $\boldsymbol{f}\big(\bar{\boldsymbol{x}} + \eta \varGamma^\xi(t), \bar{\boldsymbol{x}}_d + \eta(\bar{\boldsymbol{x}}_d^\xi - \bar{\boldsymbol{x}}_d), \boldsymbol{\theta} + \eta\xi\boldsymbol{e}^r, \boldsymbol{\delta}\big)$, and ϕ_η^ξ denotes $\phi(\mu(t|\boldsymbol{\theta} + \eta\xi\boldsymbol{e}^r) - h)$, and $|\cdot|$ denotes the Euclidian norm.

Step 2: The function $\varGamma^\xi(s)$ is of order ξ

Let $\xi \in [-a, a]$ be arbitrary. When $s_{\text{delay}} < 0$, taking the norm of both sides of (43) and applying the definition of C_3 gives

$$|\varGamma^\xi(s)|_n = \left|\int_0^s \int_0^1 \left\{\frac{\partial \boldsymbol{f}_\eta^\xi}{\partial \bar{\boldsymbol{x}}}\varGamma^\xi(t) + \frac{\partial \boldsymbol{f}_\eta^\xi}{\partial \theta_r}\xi + \frac{\partial \boldsymbol{f}_\eta^\xi}{\partial \bar{\boldsymbol{x}}_d}(\phi_\eta^\xi - \phi_\eta^0)\right\}d\eta dt\right|_n, \qquad (45)$$

where

$$\phi_\eta^\xi - \phi_\eta^0 = \phi(\mu(t|\boldsymbol{\theta} + \xi\boldsymbol{e}^r) - h) - \phi(\mu(t|\boldsymbol{\theta}) - h)$$
$$= \xi \frac{\partial \phi(\mu(t|\boldsymbol{\theta} + \eta\xi\boldsymbol{e}^r) - h)}{\partial t}\frac{\partial \mu(t|\boldsymbol{\theta} + \eta\xi\boldsymbol{e}^r)}{\partial \theta_r}, \; \eta \in [0, 1].$$

Thus, we have

$$|\varGamma^\xi(s)|_n \leq C_3|\xi| + C_3^2 T|\xi| + \int_0^s C_3|\varGamma^\xi(t)|_n dt, \; s_{\text{delay}} < 0. \qquad (46)$$

Applying Gronwall's Lemma gives

$$|\varGamma^\xi(s)|_n \leq (C_3 + C_3^2 T)\exp{(C_3 q)}|\xi|, \; s \in [0, P_1[. \qquad (47)$$

When $s_{\text{delay}} \geq 0$,

$$|\varGamma^\xi(s)|_n$$
$$= \left|\int_0^s \int_0^1 \left\{\frac{\partial \boldsymbol{f}_\eta^\xi}{\partial \bar{\boldsymbol{x}}}\varGamma^\xi(t) + \frac{\partial \boldsymbol{f}_\eta^\xi}{\partial \theta_r}\xi + \frac{\partial \boldsymbol{f}_\eta^\xi}{\partial \bar{\boldsymbol{x}}_d}\big[\varGamma^\xi(s_{\text{delay}})\right.\right.$$
$$\left.\left. + \bar{\boldsymbol{x}}(s_{\text{delay}}^\xi|\boldsymbol{\delta}, \boldsymbol{\theta}^\xi) - \bar{\boldsymbol{x}}(s_{\text{delay}}|\boldsymbol{\delta}, \boldsymbol{\theta}^\xi)\big]\right\}d\eta dt\right|_n$$
$$\leq \left|\int_0^s \int_0^1 \frac{\partial \boldsymbol{f}_\eta^\xi}{\partial \bar{\boldsymbol{x}}}\varGamma^\xi(t)d\eta dt\right|_n + \left|\int_0^s \int_0^1 \frac{\partial \boldsymbol{f}_\eta^\xi}{\partial \theta_r}\xi d\eta dt\right|_n$$
$$+ \left|\int_0^s \int_0^1 \frac{\partial \boldsymbol{f}_\eta^\xi}{\partial \bar{\boldsymbol{x}}_d}\varGamma^\xi(s_{\text{delay}})d\eta dt\right|_n$$
$$+ \left|\int_0^s \int_0^1 \frac{\partial \boldsymbol{f}_\eta^\xi}{\partial \bar{\boldsymbol{x}}_d}\big[\bar{\boldsymbol{x}}(s_{\text{delay}}^\xi|\boldsymbol{\delta}, \boldsymbol{\theta}^\xi) - \bar{\boldsymbol{x}}(s_{\text{delay}}|\boldsymbol{\delta}, \boldsymbol{\theta}^\xi)\big]d\eta dt\right|_n$$

$$
\begin{aligned}
\leq\ & (C_3 + C_3^2)\exp{(C_3 q)}|\xi| + \left|\int_{P_1}^{s}\int_0^1 \frac{\partial \boldsymbol{f}_\eta^\xi}{\partial \bar{\boldsymbol{x}}}\Gamma^\xi(t)d\eta dt\right|_n \\
& + \left|\int_{P_1}^{s}\int_0^1 \frac{\partial \boldsymbol{f}_\eta^\xi}{\partial \theta_r}\xi d\eta dt\right|_n + \left|\int_{P_1}^{s}\int_0^1 \frac{\partial \boldsymbol{f}_\eta^\xi}{\partial \bar{\boldsymbol{x}}_d}\Gamma^\xi(s_{\text{delay}})d\eta dt\right|_n \\
& + \left|\int_{P_1}^{s}\int_0^1 \frac{\partial \boldsymbol{f}_\eta^\xi}{\partial \bar{\boldsymbol{x}}_d}\big[\bar{\boldsymbol{x}}(s_{\text{delay}}^\xi|\boldsymbol{\delta},\boldsymbol{\theta}^\xi) - \bar{\boldsymbol{x}}(s_{\text{delay}}|\boldsymbol{\delta},\boldsymbol{\theta}^\xi)\big]d\eta dt\right|_n,
\end{aligned}
$$

where P_1 is a time point in the new time horizon such that

$$
\mu(P_1|\boldsymbol{\theta}) = h.
$$

Since $s_{\text{delay}} \geq 0$, it follows from the definition of s_{delay} that

$$
\left|\int_{P_1}^{s}\int_0^1 \frac{\partial \boldsymbol{f}_\eta^\xi}{\partial \bar{\boldsymbol{x}}_d}\Gamma^\xi(s_{\text{delay}})d\eta dt\right|_n \leq \int_{P_1}^{s} C_3|\Gamma^\xi(s_{\text{delay}})|_n dt \leq \int_0^{s} C_3|\Gamma^\xi(t)|_n dt,
$$

and by the mean value theorem

$$
\begin{aligned}
& \left|\int_{P_1}^{s}\int_0^1 \frac{\partial \boldsymbol{f}_\eta^\xi}{\partial \bar{\boldsymbol{x}}_z}\big[\bar{\boldsymbol{x}}(s_{\text{delay}}^\xi|\boldsymbol{\delta},\boldsymbol{\theta}^\xi) - \bar{\boldsymbol{x}}(s_{\text{delay}}|\boldsymbol{\delta},\boldsymbol{\theta}^\xi)\big]d\eta dt\right|_n \\
& \leq C_3 \int_{P_1}^{s}\int_0^1 \left|\frac{\partial \bar{\boldsymbol{x}}(z(t,\boldsymbol{\theta}+\iota\xi e^r)|\boldsymbol{\delta},\boldsymbol{\theta}+\iota\xi e^r)}{\partial z}\frac{\partial z}{\partial \theta_r}\xi\right|_n d\iota dt \leq C_3^3 q|\xi|,
\end{aligned}
\tag{48}
$$

where $\iota \in [0,1]$. Again, by applying Gronwall's Lemma, we have

$$
\begin{aligned}
|\Gamma^\xi(s)|_n &\leq (C_3 + C_3^2 T)\exp{(C_3 q)}|\xi| + C_3|\xi| + C_3^3 q|\xi| + \int_0^{s} 2C_3|\Gamma^\xi(t)|_n dt \\
&\leq (C_3 + C_3^2 T)\exp{(C_3 q)}|\xi| + (C_3 + C_3^3 q)\exp{(2C_3 q)}|\xi|.
\end{aligned}
\tag{49}
$$

Since $\xi \in [-a, a]$ is arbitrary, the function $\Gamma^\xi(s)$ is of order ξ.

Step 3: The definition of ρ and its properties

For each $\xi \in \mathbb{R}$, define the corresponding functions $\lambda^{1,\xi} : [0,q] \to \mathbb{R}^n$, $\lambda^{2,\xi} : [0,q] \to \mathbb{R}^n$ and $\lambda^{3,\xi} : [0,q] \to \mathbb{R}^n$ as follows:

$$
\begin{aligned}
\lambda^{1,\xi}(t) := \int_0^1 \Bigg\{ & \frac{\partial \boldsymbol{f}\big(\bar{\boldsymbol{x}} + \eta\Gamma^\xi(t), \bar{\boldsymbol{x}}_d + \eta(\bar{\boldsymbol{x}}_d^\xi - \bar{\boldsymbol{x}}_d), \boldsymbol{\theta} + \eta\xi e^r, \boldsymbol{\delta}\big)}{\partial \bar{\boldsymbol{x}}} \\
& - \frac{\partial \boldsymbol{f}\big(\bar{\boldsymbol{x}}, \bar{\boldsymbol{x}}_d, \boldsymbol{\theta}, \boldsymbol{\delta}\big)}{\partial \bar{\boldsymbol{x}}} \Bigg\}\Gamma^\xi(t)d\eta
\end{aligned}
\tag{50}
$$

$$
\begin{aligned}
\lambda^{2,\xi}(t) := \int_0^1 \Bigg\{ & \frac{\partial \boldsymbol{f}\big(\bar{\boldsymbol{x}} + \eta\Gamma^\xi(t), \bar{\boldsymbol{x}}_d + \eta(\bar{\boldsymbol{x}}_d^\xi - \bar{\boldsymbol{x}}_d), \boldsymbol{\theta} + \eta\xi e^r, \boldsymbol{\delta}\big)}{\partial \bar{\boldsymbol{x}}_d} \\
& - \frac{\partial \boldsymbol{f}\big(\bar{\boldsymbol{x}}, \bar{\boldsymbol{x}}_d, \boldsymbol{\theta}, \boldsymbol{\delta}\big)}{\partial \bar{\boldsymbol{x}}_d} \Bigg\}(\bar{\boldsymbol{x}}_d^\xi - \bar{\boldsymbol{x}}_d)d\eta
\end{aligned}
\tag{51}
$$

$$
\lambda^{3,\xi}(t) := \int_0^1 \left\{ \frac{\partial \boldsymbol{f}\big(\bar{\boldsymbol{x}} + \eta \Gamma^\xi(t), \bar{\boldsymbol{x}}_d + \eta(\bar{\boldsymbol{x}}_d^\xi - \bar{\boldsymbol{x}}_d), \boldsymbol{\theta} + \eta \xi \boldsymbol{e}^r, \boldsymbol{\delta}\big)}{\partial \theta_r} \right.
$$
$$
\left. - \frac{\partial \boldsymbol{f}\big(\bar{\boldsymbol{x}}, \bar{\boldsymbol{x}}_d, \boldsymbol{\theta}, \boldsymbol{\delta}\big)}{\partial \theta_r} \right\} \xi d\eta \tag{52}
$$

In addition, let the function $\rho : [-a, 0[\cup]0, a] \to \mathbb{R}$ be defined as follows:

$$
\rho(\xi) := |\xi|^{-1} \int_0^q \{ |\lambda^{1,\xi}(t)|_n + |\lambda^{2,\xi}(t)|_n + |\lambda^{3,\xi}(t)|_n \} dt \tag{53}
$$

Since the function $\Gamma^\xi(s)$ is of order ξ, it follows that

$$
\bar{\boldsymbol{x}} + \eta \Gamma^\xi(t) \to \bar{\boldsymbol{x}}, \text{ as } \xi \to 0, \tag{54}
$$

$$
\bar{\boldsymbol{x}}_d + \eta(\bar{\boldsymbol{x}}_d^\xi - \bar{\boldsymbol{x}}_d) \to \bar{\boldsymbol{x}}_d, \text{ as } \xi \to 0, \tag{55}
$$

uniformly with respective to $t \in [0, q]$ and $\eta \in [0, 1]$. Meanwhile, it is obvious that

$$
\boldsymbol{\theta} + \eta \xi \boldsymbol{e}^r \to \boldsymbol{\theta}, \text{ as } \xi \to 0, \tag{56}
$$

uniformly with respect to $\eta \in [0, 1]$. Since the convergences in (54) and (55) take place inside the ball $\mathcal{N}_n(C_1)$, the convergence in (56) takes place inside of the ball $\mathcal{N}_n(C_2)$, $\partial \boldsymbol{f}/\partial \bar{\boldsymbol{x}}$, $\partial \boldsymbol{f}/\partial \bar{\boldsymbol{x}}_z$ and $\partial \boldsymbol{f}/\partial \theta_r$ are uniformly continuous on the compact set $[0, q] \times \mathcal{N}_n(C_1) \times \mathcal{N}_n(C_1) \times \mathcal{V} \times \mathcal{N}_n(C_2)$,

$$
\frac{\partial \boldsymbol{f}\big(\bar{\boldsymbol{x}}^0 + \eta \Gamma^\xi(s), \bar{\boldsymbol{x}}_d + \eta(\bar{\boldsymbol{x}}_d^\xi - \bar{\boldsymbol{x}}_d), \boldsymbol{\theta} + \eta \xi \boldsymbol{e}^r, \boldsymbol{\delta}\big)}{\partial \bar{\boldsymbol{x}}} \to \frac{\partial \boldsymbol{f}\big(\bar{\boldsymbol{x}}^0, \bar{\boldsymbol{x}}_d, \boldsymbol{\theta}, \boldsymbol{\delta}\big)}{\partial \bar{\boldsymbol{x}}}, \text{ as } \xi \to 0,
$$

$$
\frac{\partial \boldsymbol{f}\big(\bar{\boldsymbol{x}}^0 + \eta \Gamma^\xi(s), \bar{\boldsymbol{x}}_d + \eta(\bar{\boldsymbol{x}}_d^\xi - \bar{\boldsymbol{x}}_d), \boldsymbol{\theta} + \eta \xi \boldsymbol{e}^r, \boldsymbol{\delta}\big)}{\partial \bar{\boldsymbol{x}}_d} \to \frac{\partial \boldsymbol{f}\big(\bar{\boldsymbol{x}}^0, \bar{\boldsymbol{x}}_d, \boldsymbol{\theta}, \boldsymbol{\delta}\big)}{\partial \bar{\boldsymbol{x}}_d}, \text{ as } \xi \to 0,
$$

$$
\frac{\partial \boldsymbol{f}\big(\bar{\boldsymbol{x}}^0 + \eta \Gamma^\xi(s), \bar{\boldsymbol{x}}_d + \eta(\bar{\boldsymbol{x}}_d^\xi - \bar{\boldsymbol{x}}_d), \boldsymbol{\theta} + \eta \xi \boldsymbol{e}^r, \boldsymbol{\delta}\big)}{\partial \theta_r} \to \frac{\partial \boldsymbol{f}\big(\bar{\boldsymbol{x}}^0, \bar{\boldsymbol{x}}_d, \boldsymbol{\theta}, \boldsymbol{\delta}\big)}{\partial \theta_r}, \text{ as } \xi \to 0,
$$

$$
\frac{\partial \bar{\boldsymbol{x}}(s_{\text{delay}}^{\xi,\iota}|\boldsymbol{\delta}, \boldsymbol{\theta} + \iota \xi \boldsymbol{e}^r)}{\partial s_{\text{delay}}} \to \frac{\partial \bar{\boldsymbol{x}}(s_{\text{delay}}|\boldsymbol{\delta}, \boldsymbol{\theta})}{\partial s_{\text{delay}}}, \text{ as } \xi \to 0,
$$

$$
\frac{\partial s_{\text{delay}}^{\xi,\iota}}{\partial \theta_r} \to \frac{\partial s_{\text{delay}}}{\partial \theta_r}, \text{ as } \xi \to 0,
$$

uniformly with respect to $t \in [0, q]$, $\eta \in [0, 1]$ and $\iota \in [0, 1]$, where $s_{\text{delay}}^{\xi,\iota}$ is the corresponding delayed time of the control $\boldsymbol{\theta} + \iota \xi \boldsymbol{e}^r$. These results together with (49) indicate that $|\xi|^{-1} \lambda^{1,\xi} \to \boldsymbol{0}$, $|\xi|^{-1} \lambda^{2,\xi} \to \boldsymbol{0}$ and $|\xi|^{-1} \lambda^{3,\xi} \to \boldsymbol{0}$ uniformly on $[0, q]$ as $\xi \to 0$. Thus,

$$
\lim_{\xi \to 0} \rho(\xi) = \boldsymbol{0}. \tag{57}
$$

Step 4: The final step

Let $\xi \in [-a, 0[\cup]0, a]$ be arbitrary but fixed. Then, it follows from (43) that

$$
\begin{aligned}
\Gamma^{\xi}(s) = & \int_0^s \left[\lambda^{1,\xi}(t) + \lambda^{2,\xi}(t) + \lambda^{3,\xi}(t)\right] dt + \int_0^s \frac{\partial \boldsymbol{f}\left(\bar{\boldsymbol{x}}, \bar{\boldsymbol{x}}_d, \boldsymbol{\theta}, \boldsymbol{\delta}\right)}{\partial \bar{\boldsymbol{x}}} \Gamma^{\xi}(t) dt \\
& + \int_0^s \frac{\partial \boldsymbol{f}\left(\bar{\boldsymbol{x}}, \bar{\boldsymbol{x}}_d, \boldsymbol{\theta}, \boldsymbol{\delta}\right)}{\partial \bar{\boldsymbol{x}}_d} [\Gamma^{\xi}(s_{\text{delay}} + \frac{\partial \bar{\boldsymbol{x}}(s_{\text{delay}}^{\xi,\iota}|(\boldsymbol{\delta}, \boldsymbol{\theta} + \iota\xi\boldsymbol{e}^r))}{\partial s_{\text{delay}}} \frac{\partial s_{\text{delay}}^{\xi,\iota}}{\partial \theta_r} \xi] dt \\
& + \int_0^s \frac{\partial \boldsymbol{f}\left(\bar{\boldsymbol{x}}, \bar{\boldsymbol{x}}_d, \boldsymbol{\theta}, \boldsymbol{\delta}\right)}{\partial \theta_r} \xi dt
\end{aligned}
$$

$$(58)$$

Furthermore, integrating the auxiliary system gives

$$
\begin{aligned}
\Lambda^r(s|\boldsymbol{\delta}, \boldsymbol{\theta}) = & \int_0^s \frac{\partial \boldsymbol{f}\left(\bar{\boldsymbol{x}}, \bar{\boldsymbol{x}}_d, \boldsymbol{\theta}, \boldsymbol{\delta}\right)}{\partial \bar{\boldsymbol{x}}} \Lambda^r(t|\boldsymbol{\delta}, \boldsymbol{\theta}) dt \\
& + \int_0^s \frac{\partial \boldsymbol{f}\left(\bar{\boldsymbol{x}}, \bar{\boldsymbol{x}}_d, \boldsymbol{\theta}, \boldsymbol{\delta}\right)}{\partial \bar{\boldsymbol{x}}_d} \left[\Lambda^r(s_{\text{delay}}|\boldsymbol{\delta}, \boldsymbol{\theta}) + \frac{\partial \bar{\boldsymbol{x}}(s_{\text{delay}}|(\boldsymbol{\delta}, \boldsymbol{\theta}))}{\partial s_{\text{delay}}} \frac{\partial s_{\text{delay}}}{\partial \theta_r}\right] dt \\
& + \int_0^s \frac{\partial \boldsymbol{f}\left(\bar{\boldsymbol{x}}, \bar{\boldsymbol{x}}_d, \boldsymbol{\theta}, \boldsymbol{\delta}\right)}{\partial \theta_r} dt
\end{aligned}
$$

$$(59)$$

Multiplying (58) by ξ^{-1}, and subtracting it from (59) yields

$$
\begin{aligned}
& \xi^{-1} \Gamma^{\xi}(s) - \Lambda^r(s|\boldsymbol{\delta}, \boldsymbol{\theta}) \\
= & \xi^{-1} \int_0^s \left[\lambda^{1,\xi}(t) + \lambda^{2,\xi}(t) + \lambda^{3,\xi}(t)\right] dt \\
& + \int_0^s \frac{\partial \boldsymbol{f}\left(\bar{\boldsymbol{x}}, \bar{\boldsymbol{x}}_d, \boldsymbol{\theta}, \boldsymbol{\delta}\right)}{\partial \bar{\boldsymbol{x}}} (\xi^{-1} \Gamma^{\xi}(t) - \Lambda^r(s|\boldsymbol{\delta}, \boldsymbol{\theta})) dt \\
& + \int_0^s \frac{\partial \boldsymbol{f}\left(\bar{\boldsymbol{x}}, \bar{\boldsymbol{x}}_d, \boldsymbol{\theta}, \boldsymbol{\delta}\right)}{\partial \bar{\boldsymbol{x}}_z} \left[(\xi^{-1} \Gamma^{\xi}(s_{\text{delay}}) - \Lambda^r(s_{\text{delay}}|\boldsymbol{\delta}, \boldsymbol{\theta}))\right. \\
& + \left. \frac{\partial \bar{\boldsymbol{x}}(s_{\text{delay}}^{\xi,\iota}|(\boldsymbol{\delta}, \boldsymbol{\theta} + \iota\xi\boldsymbol{e}^r))}{\partial s_{\text{delay}}} \frac{\partial s_{\text{delay}}^{\xi,\iota}}{\partial \theta_r} - \frac{\partial \bar{\boldsymbol{x}}(s_{\text{delay}}|(\boldsymbol{\delta}, \boldsymbol{\theta}))}{\partial s_{\text{delay}}} \frac{\partial s_{\text{delay}}}{\partial \theta_r}\right] dt
\end{aligned}
$$

$$(60)$$

Let

$$
\begin{aligned}
\bar{\rho}(\xi) = \rho(\xi) + \int_0^s C_3 \Bigg| & \frac{\partial \bar{\boldsymbol{x}}(s_{\text{delay}}^{\xi,\iota}|(\boldsymbol{\delta}, \boldsymbol{\theta} + \iota\xi\boldsymbol{e}^r))}{\partial s_{\text{delay}}} \frac{\partial s_{\text{delay}}^{\xi,\iota}}{\partial \theta_r} \\
& - \frac{\partial \bar{\boldsymbol{x}}(s_{\text{delay}}|(\boldsymbol{\delta}, \boldsymbol{\theta}))}{\partial s_{\text{delay}}} \frac{\partial s_{\text{delay}}}{\partial \theta_r}\Bigg|_n dt.
\end{aligned}
$$

Then, it is easy to see that $\rho'(\xi) \to 0$, as $\xi \to 0$. Therefore,

$$|\xi^{-1}\Gamma^{\xi}(s) - \Lambda^r(s|\boldsymbol{\delta}, \boldsymbol{\theta})|_n$$

$$\leq \rho(\xi) + \int_0^s C_3 |\xi^{-1}\Gamma^{\xi}(t) - \Lambda^r(t|\boldsymbol{\delta}, \boldsymbol{\theta})|_n dt$$

$$+ \int_0^s C_3 |\xi^{-1}\Gamma^{\xi}(s_{\mathrm{delay}}) - \Lambda^r(s_{\mathrm{delay}}|\boldsymbol{\delta}, \boldsymbol{\theta})|_n dt$$

$$+ \int_0^s C_3 \left| \frac{\partial \bar{\boldsymbol{x}}(s_{\mathrm{delay}}^{\xi,\iota}|(\boldsymbol{\delta}, \boldsymbol{\theta} + \iota\xi\boldsymbol{e}^r)}{\partial s_{\mathrm{delay}}} \frac{\partial s_{\mathrm{delay}}^{\xi,\iota}}{\partial \theta_r} - \frac{\partial \bar{\boldsymbol{x}}(s_{\mathrm{delay}}|(\boldsymbol{\delta}, \boldsymbol{\theta}))}{\partial s_{\mathrm{delay}}} \frac{\partial s_{\mathrm{delay}}}{\partial \theta_r} \right|_n dt$$

$$\leq \bar{\rho}(\xi) + \int_0^s 2C_3 |\xi^{-1}\Gamma^{\xi}(t) - \Lambda^r(t|\boldsymbol{\delta}, \boldsymbol{\theta})|_n dt$$

By Gronwall's Lemma

$$|\xi^{-1}\Gamma^{\xi}(s) - \Lambda^r(s|\boldsymbol{\delta}, \boldsymbol{\theta})|_n \leq \bar{\rho}(\xi) \exp(2C_3 q), \ \ s \in [0, q]. \tag{61}$$

Noting that $\xi \in [-a, 0[\cup]0, a]$ is arbitrary, we can take the limit as $\xi \to 0$ in (61) and then apply (57) to get

$$\lim_{\xi \to 0} \xi^{-1}\Gamma^{\xi}(s) = \Lambda^r(s|\boldsymbol{\delta}, \boldsymbol{\theta}), \ \ s \in [0, q].$$

References

1. Stengel, R.F., Ghigliazza, R., Kulkarni, N., Laplace, O.: Optimal control of innate immune response. Optimal Control Applications and Methods. 23, 91-104 (2002)
2. Denis-Vidal, L., Jauberthie, C., Joly-Blanchard, G.:Identifiability of a nonlinear delayed-differential aerospace model. IEEE Transactions on Automatic Control. 51, 154-158 (2006)
3. Chai, Q.Q., Loxton, R., Teo, K.L., Yang, C.H.: A class of optimal state-delay control problems. Nonlinear Analysis: Real World Applications. 14, 1536-1550 (2013)
4. Chai, Q.Q., Loxton, R., Teo, K.L., Yang, C.H.: A unified parameter identification method for nonlinear time-delay systems. Journal of Industrial and Management Optimization. 9, 471-486 (2013)
5. Göllmann, L., Kern, D., Maurer, H.:Optimal control problems with delays in state and control variables subject to mixed control-state constraints. Optimal Control Applications and Methods. 30, 341-365 (2009)
6. Manu, M.Z., Mohammad, J.:Time-Delay Systems: Analysis, Optimization and Applications. Elsevier Science Inc., New York (1987)
7. Göllmann, L., Maurer, H.: Theory and applications of optimal control problems with multiple time-delays. Journal of Industrial and Management Optimization. 10, 413-441 (2014)
8. Betts, J.T., Campbell, S.L., Thompson, K.C.: Optimal control software for constrained nonlinear systems with delays. In Proceedings, IEEE Multi Conference on Systems and Control. 444-449 (2011)
9. Betts, J.T., Campbell, S.L., Thompson, K.: Optimal control of delay partial differential equations. In: Control and optimization with differential-algebraic constraints. 213231. SIAM, Philadelphia (2012)

10. Betts, J.T., Campbell, S.L., Thompson, K.: Direct transcription solution of optimal control problems with control delays, AIP Conference Proceedings 1389, 38 (2011); doi: 10.1063/1.3636665

11. Deshmuk, V. Ma, H., Butcher, E.: Optimal control of parametrically excited linear delay differential systems via chebyshev polynomials, Optimal Control Applications and Methods. 27(3), 123-136 (2006)

12. Teo, K.L., Goh, C.J., Wong, K.H.: A Unified Computational Approach to Optimal Control Problems. Longman Scientific and Technical, Essex (1991)

13. Chai, Q.Q., Yang, C.H., Teo, K.L., Gui, W.H.:Time-delay optimal control of an industrial-scale evaporation process sodium aluminate solution. Control Engineering Practice. 20, 618-628 (2012)

14. Wong, K.H., Jennings, L.S., Benyah, F.: The control parameterization enhancing transform for constrained time-delay optimal control problems. ANZIAM Journal. 43, E154-E185 (2001/02)

15. Nocedal, J., Wright, S.J.: Numerical Optimization, 2nd Edition, Springer, New York (2006)

16. Lin, Q., Loxton, R., Teo, k.l.: The control parameterization method for nonlinear optimal control: A survey. Journal of Industrial and Management Optimization. 10, 275-309 (2014)

17. Lee, H.W.J., Teo, K.L., Rehbock, V., Jennings, L.S.: Control parameterization enhancing technique for optimal discrete-valued control problems. Automatica. 35, 1401-1407 (1999)

18. Lee, H.W.J., Teo, K.L., Rehbock, V., Jennings, L.S.: Control parameterization enhancing technique for time optimal control problems. Dynamic Systems and Applications. 6, 243-262 (1997)

19. Liu, C., Gong, z., Shen, B., Feng, E.: Modelling and optimal control for a fed-batch fermentation process. Applied Mathematical Modelling. 37, 695-706 (2013)

20. Blanchard, E., Loxton, R., Rehbock, V.: Optimal control of impulsive switched systems with minimum subsystem durations. Journal of Global Optimization. 4, 737-750 (2014)

21. Yu, C.J., Li, B., Loxton, R., Teo, K.L.: Optimal discrete-valued control computation. Journal of Global Optimization, 56, 503-518 (2013)