

NOTICE: This is the author's version of a work that was accepted for publication in Applied Mathematics and Computations. Changes resulting from the publishing process, such as peer review, editing, corrections, structural formatting, and other quality control mechanisms may not be reflected in this document. Changes may have been made to this work since it was submitted for publication. A definitive version was subsequently published Applied Mathematics and Computations, Vol. 219, Issue 7. (2013). doi: 10.1016/j.amc.2013.03.015

Time-delay estimation for nonlinear systems with piecewise-constant input

Qinqin Chai^{1,2*}, Ryan Loxton², Kok Lay Teo², Chunhua Yang¹

¹ School of Information Science & Engineering, Central South University, Changsha, China

² Department of Mathematics & Statistics, Curtin University, Perth, Australia

* Corresponding author: kppqing@163.com

Abstract

We consider a general nonlinear time-delay system in which the input signal is piecewise-constant. Such systems arise in a wide range of industrial applications, including evaporation and purification processes and chromatography. We assume that the time-delays—one involving the state variables and the other involving the input variables—are unknown and need to be estimated using experimental data. We formulate the problem of estimating the unknown delays as a nonlinear optimization problem in which the cost function measures the least-squares error between predicted and measured system output. The main difficulty with this problem is that the delays are decision variables to be optimized, rather than fixed values. Thus, conventional optimization techniques are not directly applicable. We propose a new computational approach based on a novel derivation of the cost function's gradient. We then apply this approach to estimate the time-delays in two industrial chemical processes: a zinc sulphate purification process and a sodium aluminate evaporation process.

Keywords: parameter identification; time-delay system; nonlinear optimization; purification processes; evaporation processes

1 Introduction

Time-delays are inherent in many industrial processes, including evaporation processes [1], chromatography processes [2], distillation processes [3], and purification processes [4]. Such processes can be controlled by varying certain input variables—for example, flow rates, temperatures, and pressures. If the time-delays are known, then the problem of determining the optimal input variables (as functions of time) so that the total system cost is minimized is a type of *optimal control problem*. Such problems can be solved numerically using well-known computational techniques [4, 5, 6, 7].

In many situations, however, the time-delays are not known exactly. In this case, the delays first need to be estimated before optimal control techniques can be applied. Thus,

delay estimation is a crucial issue and has attracted significant research attention over the past decade [8]. The vast majority of delay estimation methods are only applicable to simple systems with linear dynamics and a single delay [9, 10, 11, 12]. One of the few methods available for handling general nonlinear systems with multiple time-delays is the optimization-based approach developed in [13]. In this approach, the problem of estimating the time-delays is formulated as a dynamic optimization problem in which the cost function measures the discrepancy between predicted and observed system output at a set of sample times. Solving this optimization problem yields the delay values that best fit the given experimental data.

The optimization-based approach developed in [13] is designed for systems with state-delays rather than input-delays. For systems with input-delays, if the input function is smooth, then the system dynamics will be continuously differentiable with respect to the input-delays, and thus the approach proposed in [13] can be easily modified to estimate the input-delays in this case. Unfortunately, the input function is often non-smooth in practical applications. For example, in the evaporation process described in [1], the input function represents the steam flow rate, which must be kept constant and is only changed at 5 minute intervals to ensure process stability. Moreover, the input function represents the solution flow rate, which changes in piecewise manner to match the scheduling of the production process. Thus, the steam flow rate is a non-smooth piecewise-constant input function. The estimation method in [13] is not applicable in such situations.

In this paper, we consider the time-delay estimation problem for nonlinear systems in which the input function is piecewise-constant. Such estimation problems arise, for example, in evaporation and purification processes [1, 4]. We assume that the system under consideration contains one state-delay and one input-delay, both of which are unknown and need to be estimated using experimental data. Since the input function is discontinuous, the estimation method in [13] is not applicable in this case. The purpose of this paper is to develop a new method for estimating the time-delays. As with [13], we formulate the delay estimation problem as a dynamic optimization problem in which the cost function measures the least-squares error between predicted and observed system output. The main focus of the paper is on the derivation of a computational procedure for determining the gradient of the cost function. This procedure, which involves integrating an auxiliary impulsive system with instantaneous jumps forward in time, is far more complex than the procedure given in [13], which does not involve any jumps. Moreover, because of the discontinuous nature of the input function, the cost function's gradient does not exist at certain points. We propose a heuristic strategy for dealing this complication. This heuristic strategy can be combined with our gradient computation procedure to solve the estimation problem using standard nonlinear programming algorithms. We finally conclude the paper by showing that this approach can successfully estimate the time-delays in two large-scale chemical engineering systems.

2 Problem formulation

Consider the following nonlinear time-delay system:

$$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), \mathbf{x}(t - \alpha), \mathbf{u}(t), \mathbf{u}(t - \beta)), \quad t \in [0, T], \quad (1)$$

$$\mathbf{x}(t) = \boldsymbol{\phi}(t), \quad t \leq 0, \quad (2)$$

where $T > 0$ is a given *terminal time*; $\mathbf{x}(t) = [x_1(t), \dots, x_n(t)]^\top \in \mathbb{R}^n$ is the *state vector*; $\mathbf{u}(t) = [u_1(t), \dots, u_r(t)]^\top \in \mathbb{R}^r$ is the *input vector*; α and β are unknown time-delays that need to be determined; and $\mathbf{f} : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^r \times \mathbb{R}^r \rightarrow \mathbb{R}^n$ and $\boldsymbol{\phi} : \mathbb{R} \rightarrow \mathbb{R}^n$ are given functions. Many dynamic processes in chemical engineering—for example, the distillation process described in [14]—can be modeled by equations (1) and (2). We assume that \mathbf{f} , \mathbf{g} , and $\boldsymbol{\phi}$ are continuously differentiable. We also assume that there exists a positive real number $L_1 > 0$ such that for all $\mathbf{x}', \mathbf{x}'' \in \mathbb{R}^n$ and $\mathbf{u}', \mathbf{u}'' \in \mathbb{R}^r$,

$$|\mathbf{f}(\mathbf{x}', \mathbf{x}'', \mathbf{u}', \mathbf{u}'')| \leq L_1(1 + |\mathbf{x}'| + |\mathbf{x}''| + |\mathbf{u}'| + |\mathbf{u}''|), \quad (3)$$

where $|\cdot|$ denotes the Euclidean norm. This assumption is standard in the control systems literature [4, 15, 16, 17].

The output $\mathbf{y}(t)$ of system (1)-(2) is defined by

$$\mathbf{y}(t) = \mathbf{g}(\mathbf{x}(t), \mathbf{x}(t - \alpha)), \quad t \in [0, T], \quad (4)$$

where $\mathbf{g} : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^q$ is a given continuously differentiable function.

We refer to α as the *state-delay* and β as the *input-delay*. The exact values of these delays are unknown; the only information we are given is that α lies within the interval $[\alpha_{\min}, \alpha_{\max}]$ and β lies within the interval $[\beta_{\min}, \beta_{\max}]$, where $\alpha_{\min} \leq 0$ and $\beta_{\min} > 0$. Thus, we have the following bound constraints:

$$\alpha_{\min} \leq \alpha \leq \alpha_{\max}, \quad (5)$$

$$\beta_{\min} \leq \beta \leq \beta_{\max}. \quad (6)$$

We assume that the input signal \mathbf{u} is a given piecewise-constant function (this is the case in many engineering systems). Hence, \mathbf{u} can be expressed as follows:

$$\mathbf{u}(t) = \boldsymbol{\sigma}^i, \quad t \in [t_{i-1}, t_i), \quad i = 1, \dots, p, \quad (7)$$

where $\boldsymbol{\sigma}^i \in \mathbb{R}^r$, $i = 1, \dots, p$, are given vectors and t_i , $i = 0, \dots, p$, are given time points such that $-\beta_{\max} = t_0 < t_1 < \dots < t_p = T$. Equation (7) can be rewritten as

$$\mathbf{u}(t) = \sum_{i=1}^p \boldsymbol{\sigma}^i \chi_{[t_{i-1}, t_i)}(t), \quad t \in [-\beta_{\max}, T], \quad (8)$$

where the characteristic function $\chi_{[t_{i-1}, t_i)} : \mathbb{R} \rightarrow \mathbb{R}$ is defined by

$$\chi_{[t_{i-1}, t_i)}(t) = \begin{cases} 1, & \text{if } t \in [t_{i-1}, t_i), \\ 0, & \text{otherwise.} \end{cases}$$

For each pair $(\alpha, \beta) \in [\alpha_{\min}, \alpha_{\max}] \times [\beta_{\min}, \beta_{\max}]$, let $\mathbf{x}(\cdot|\alpha, \beta)$ denote the corresponding solution of system (1)-(2). Substituting $\mathbf{x}(\cdot|\alpha, \beta)$ into equation (4) gives $\mathbf{y}(\cdot|\alpha, \beta)$, the predicted system output corresponding to (α, β) . Mathematically,

$$\mathbf{y}(t|\alpha, \beta) = \mathbf{g}(\mathbf{x}(t|\alpha, \beta), \mathbf{x}(t - \alpha|\alpha, \beta)), \quad t \leq T. \quad (9)$$

Suppose that the output from system (1)-(2) has been measured experimentally at a set of sample times $\{\tau_j\}_{j=1}^m \subset [0, T]$. Let $\hat{\mathbf{y}}^j \in \mathbb{R}^q$ denote the measured output at the j th sample time. Our goal is to use the experimental data $\{(\tau_j, \hat{\mathbf{y}}^j)\}_{j=1}^m$ to identify the time-delays α and β . We formulate this problem mathematically as follows.

Problem (P). Choose the state-delay α and the input-delay β to minimize the cost function

$$J(\alpha, \beta) = \sum_{j=1}^m |\mathbf{y}(\tau_j|\alpha, \beta) - \hat{\mathbf{y}}^j|^2 \quad (10)$$

subject to the dynamic system (1)-(2) and the bound constraints (5)-(6).

Problem (P) is a dynamic optimization problem governed by the time-delay system (1)-(2). The most interesting aspect of Problem (P) is that the time-delays in (1)-(2) are actually decision variables to be chosen optimally. This is highly unusual; in most optimization problems involving time-delay systems, the delays are fixed and known, and the control input function is the decision variable to be chosen optimally [1, 4, 8]. In Problem (P), the input function is known, and the delays are the variables that need to be optimized.

We now conclude this section by showing that Problem (P) can be transformed into a *switched system optimal control problem*.

First, from (8),

$$\mathbf{u}(t - \beta) = \sum_{i=1}^p \sigma^i \chi_{[t_{i-1}, t_i)}(t - \beta) = \sum_{i=1}^p \sigma^i \chi_{[t_{i-1} + \beta, t_i + \beta)}(t) = \sum_{i=1}^p \sigma^i \chi_{[v_{i-1}, v_i)}(t), \quad (11)$$

where v_i , $i = 0, \dots, p$ are new decision variables defined by

$$v_i = t_i + \beta, \quad i = 0, \dots, p. \quad (12)$$

It follows from (12) that

$$v_i - t_i = v_{i-1} - t_{i-1}, \quad i = 1, \dots, p. \quad (13)$$

Substituting (11) into (1) gives

$$\dot{\mathbf{x}}(t) = \mathbf{f}^i(\mathbf{x}(t), \mathbf{x}(t - \alpha), \mathbf{u}(t)), \quad t \in [v_{i-1}, v_i) \cap [0, T], \quad i = 1, \dots, p, \quad (14)$$

where

$$\mathbf{f}^i(\mathbf{x}(t), \mathbf{x}(t - \alpha), \mathbf{u}(t)) = \mathbf{f}(\mathbf{x}(t), \mathbf{x}(t - \alpha), \mathbf{u}(t), \sigma^i).$$

System (14) is a *switched system* in which the dynamics change instantaneously at the *switching times* v_i , $i = 1, \dots, p$.

Problem (P) can now be restated as follows: Choose the state-delay α and the switching times $v_i, i = 1, \dots, p$ to minimize (10) subject to the switched system (14), the initial condition (2), and the constraints (5)-(6) and (13). This is an example of a switched system optimal control problem. Such problems have been the subject of active research over the last decade (see, for example, [18, 19, 20, 21, 22] and the references cited therein). In particular, the well-known *time-scaling transformation* is a powerful tool for solving switched system optimal control problems (see [15, 16, 23, 24]). Unfortunately, the time-scaling transformation is not applicable to time-delay systems such as system (14) defined above. Thus, a new method is needed to solve Problem (P).

3 State variation

Our goal is to solve Problem (P) using nonlinear optimization techniques. To do this, we need the partial derivatives of J with respect to the decision variables α and β . However, since J is not an explicit function of α and β , these partial derivatives cannot be determined using standard differentiation rules. To derive formulae for the partial derivatives of J , we first need to consider the *state variation* with respect to α and β .

3.1 State variation with respect to the state-delay

Define

$$\psi(t) = \begin{cases} \dot{\phi}(t), & \text{if } t \leq 0, \\ \mathbf{f}(\mathbf{x}(t), \mathbf{x}(t - \alpha), \mathbf{u}(t), \mathbf{u}(t - \beta)), & \text{if } t \in (0, T]. \end{cases}$$

Furthermore, we let $\frac{\partial}{\partial \tilde{\mathbf{x}}}$ denote differentiation with respect to the delayed state argument. We will use this notation frequently throughout the paper.

The solution $\mathbf{x}(\cdot|\alpha, \beta)$ of system (1)-(2) is normally viewed as a function of time, with α and β being fixed values. By instead of fixing $t \in (-\infty, T]$ while allowing α and β to vary, we obtain the function $\mathbf{x}(t|\cdot, \cdot) : [\alpha_{\min}, \alpha_{\max}] \times [\beta_{\min}, \beta_{\max}] \rightarrow \mathbb{R}^n$ whose value at (α, β) is $\mathbf{x}(t|\alpha, \beta)$. The partial derivative of $\mathbf{x}(t|\cdot, \cdot)$ with respect to α is called the *state variation* with respect to α . The following result, which can be proved in a similar manner to the main proof in [13], gives a method for determining this state variation.

Theorem 1. *Let $t \in (0, T]$ be a fixed time point. Then $\mathbf{x}(t|\cdot, \cdot)$ is differentiable with respect to the state-delay α . In fact, for each $(\alpha, \beta) \in [\alpha_{\min}, \alpha_{\max}] \times [\beta_{\min}, \beta_{\max}]$,*

$$\frac{\partial \mathbf{x}(t|\alpha, \beta)}{\partial \alpha} = \mathbf{\Lambda}(t|\alpha, \beta), \quad (15)$$

where $\mathbf{\Lambda}(\cdot|\alpha, \beta)$ satisfies the auxiliary time-delay system

$$\begin{aligned} \dot{\mathbf{\Lambda}}(t) = & \frac{\partial \mathbf{f}(\mathbf{x}(t), \mathbf{x}(t - \alpha), \mathbf{u}(t), \mathbf{u}(t - \beta))}{\partial \mathbf{x}} \mathbf{\Lambda}(t) \\ & + \frac{\partial \mathbf{f}(\mathbf{x}(t), \mathbf{x}(t - \alpha), \mathbf{u}(t), \mathbf{u}(t - \beta))}{\partial \tilde{\mathbf{x}}} \mathbf{\Lambda}(t - \alpha) \\ & - \frac{\partial \mathbf{f}(\mathbf{x}(t), \mathbf{x}(t - \alpha), \mathbf{u}(t), \mathbf{u}(t - \beta))}{\partial \tilde{\mathbf{x}}} \psi(t - \alpha) \end{aligned} \quad (16)$$

with initial condition

$$\mathbf{\Lambda}(t) = \mathbf{0}, \quad t \leq 0. \quad (17)$$

According to Theorem 1, the state variation with respect to α can be computed by solving the auxiliary time-delay system (16)-(17). This result is a simple extension of the main result in [13], which pertains to systems with multiple state-delays but no input delays. To solve Problem (P), we also need the state variation with respect to β . Unfortunately, the results in [13], which are based on the assumption that the system dynamics are continuous with respect to the time-delays, cannot be used to determine this state variation. Indeed, since the input function \mathbf{u} is discontinuous, the dynamics (1) are clearly discontinuous with respect to β . In the next subsection, we describe a new method for computing the state variation with respect to β .

3.2 State variation with respect to the input-delay

3.2.1 Preliminaries

Before deriving the state variation with respect to β , we first need to derive several preliminary results. Let $(\alpha, \beta) \in [\alpha_{\min}, \alpha_{\max}] \times [\beta_{\min}, \beta_{\max}]$ be a fixed pair. Define

$$\mathcal{S} = [\beta_{\min} - \beta, \beta_{\max} - \beta].$$

Note that \mathcal{S} is a non-empty closed interval of positive measure. Clearly,

$$\epsilon \in \mathcal{S} \quad \iff \quad \beta + \epsilon \in [\beta_{\min}, \beta_{\max}].$$

Now, for each $\epsilon \in \mathcal{S}$, define

$$\varphi^\epsilon(t) = \mathbf{x}(t|\alpha, \beta + \epsilon) - \mathbf{x}(t|\alpha, \beta), \quad t \leq T. \quad (18)$$

By (2),

$$\varphi^\epsilon(t) = \mathbf{0}, \quad t \leq 0. \quad (19)$$

Since the system dynamics satisfy the linear growth condition (3), it can be shown (see [13]) that there exists a positive real number $L_2 > 0$ such that

$$|\mathbf{x}(t|\alpha, \beta + \epsilon)| \leq L_2, \quad t \in [-\alpha_{\max}, T], \quad \epsilon \in \mathcal{S}. \quad (20)$$

Our first preliminary result is stated and proved below.

Lemma 1. *There exists a positive real number $L_3 > 0$ such that for all $\epsilon \in \mathcal{S}$ of sufficiently small magnitude,*

$$|\varphi^\epsilon(t)| \leq L_3|\epsilon|, \quad t \in (-\infty, T]. \quad (21)$$

Proof. Let $\epsilon \in \mathcal{S}$ be such that

$$|\epsilon| < \frac{1}{2} \min \{t_i - t_{i-1}\}_{i=1}^p.$$

For each $i = 1, \dots, p$, define

$$I_i = (t_{i-1} + |\epsilon|, t_i - |\epsilon|).$$

Furthermore, for each $i = 0, \dots, p$, define

$$J_i = \begin{cases} [t_0, t_0 + |\epsilon|], & i = 0, \\ [t_i - |\epsilon|, t_i + |\epsilon|], & i = 1, \dots, p-1, \\ [t_p - |\epsilon|, t_p], & i = p. \end{cases}$$

Note that $\{I_i\}_{i=1}^p$ and $\{J_i\}_{i=0}^p$ form a partition of $[-\beta_{\max}, T]$. Also,

$$|J_i| \leq 2|\epsilon|, \quad i = 0, \dots, p,$$

and

$$\mathbf{u}(s) = \mathbf{u}(s - \epsilon) = \boldsymbol{\sigma}^i, \quad s \in I_i, \quad i = 1, \dots, p. \quad (22)$$

Now, if $t \leq 0$, then $\boldsymbol{\varphi}^\epsilon(t) = \mathbf{0}$ and the proof is complete. Thus, assume that $t > 0$. Then

$$\begin{aligned} |\boldsymbol{\varphi}^\epsilon(t)| &\leq |\mathbf{x}^\epsilon(t) - \mathbf{x}(t)| \\ &\leq \int_0^t \left| \mathbf{f}(\mathbf{x}^\epsilon(s), \mathbf{x}^\epsilon(s - \alpha), \mathbf{u}(s), \mathbf{u}(s - \beta - \epsilon)) - \mathbf{f}(\mathbf{x}(s), \mathbf{x}(s - \alpha), \mathbf{u}(s), \mathbf{u}(s - \beta)) \right| ds, \end{aligned}$$

where $\mathbf{x}^\epsilon(s) = \mathbf{x}(s|\alpha, \beta + \epsilon)$ and $\mathbf{x}(s) = \mathbf{x}(s|\alpha, \beta)$.

Thus, since \mathbf{x}^ϵ is uniformly bounded with respect to $\epsilon \in \mathcal{S}$ (recall (20)) and \mathbf{f} is continuously differentiable, there exists a constant $M_1 > 0$ such that

$$|\boldsymbol{\varphi}^\epsilon(t)| \leq M_1 \int_0^t |\boldsymbol{\varphi}^\epsilon(s)| ds + M_1 \int_0^t |\boldsymbol{\varphi}^\epsilon(s - \alpha)| ds + M_1 \int_0^t |\mathbf{u}(s - \beta - \epsilon) - \mathbf{u}(s - \beta)| ds.$$

By shifting the time variable in the second and third integrals and then using (19), we obtain

$$\begin{aligned} |\boldsymbol{\varphi}^\epsilon(t)| &\leq M_1 \int_0^t |\boldsymbol{\varphi}^\epsilon(s)| ds + M_1 \int_{-\alpha}^{t-\alpha} |\boldsymbol{\varphi}^\epsilon(s)| ds + M_1 \int_{-\beta}^{t-\beta} |\mathbf{u}(s - \epsilon) - \mathbf{u}(s)| ds \\ &\leq 2M_1 \int_0^t |\boldsymbol{\varphi}^\epsilon(s)| ds + M_1 \int_{-\beta}^{t-\beta} |\mathbf{u}(s - \epsilon) - \mathbf{u}(s)| ds \\ &= 2M_1 \int_0^t |\boldsymbol{\varphi}^\epsilon(s)| ds + M_1 \sum_{i=1}^p \int_{I_i \cap (-\beta, t-\beta)} |\mathbf{u}(s - \epsilon) - \mathbf{u}(s)| ds \\ &\quad + M_1 \sum_{i=0}^p \int_{J_i \cap (-\beta, t-\beta)} |\mathbf{u}(s - \epsilon) - \mathbf{u}(s)| ds. \end{aligned}$$

Hence, by (22),

$$|\boldsymbol{\varphi}^\epsilon(t)| \leq 2M_1 \int_0^t |\boldsymbol{\varphi}^\epsilon(s)| ds + M_1 M_2 \sum_{i=0}^p |J_i|,$$

where $M_2 = \max_{j \neq k} |\boldsymbol{\sigma}^j - \boldsymbol{\sigma}^k|$. Since $|J_i| \leq 2|\epsilon|$, we have

$$|\boldsymbol{\varphi}^\epsilon(t)| \leq 2M_1 \int_0^t |\boldsymbol{\varphi}^\epsilon(s)| ds + 2(p+1)M_1M_2|\epsilon|.$$

Finally, applying the Gronwall-Bellman Lemma [17] yields

$$|\boldsymbol{\varphi}^\epsilon(t)| \leq 2(p+1)M_1M_2 \exp\{2M_1T\}|\epsilon|.$$

This completes the proof. \square

For each $\epsilon \in \mathcal{S}$, define

$$\bar{\mathbf{f}}^\epsilon(s, \eta, \boldsymbol{\sigma}) = \mathbf{f}(\mathbf{x}(s) + \eta\boldsymbol{\varphi}^\epsilon(s), \mathbf{x}(s - \alpha) + \eta\boldsymbol{\varphi}^\epsilon(s - \alpha), \mathbf{u}(s), \boldsymbol{\sigma}),$$

where, as in the proof of Lemma 1, $\mathbf{x}(t) = \mathbf{x}(t|\alpha, \beta)$. Then by the chain rule,

$$\frac{\partial \bar{\mathbf{f}}^\epsilon(s, \eta, \boldsymbol{\sigma})}{\partial \eta} = \frac{\partial \bar{\mathbf{f}}^\epsilon(s, \eta, \boldsymbol{\sigma})}{\partial \mathbf{x}} \boldsymbol{\varphi}^\epsilon(s) + \frac{\partial \bar{\mathbf{f}}^\epsilon(s, \eta, \boldsymbol{\sigma})}{\partial \tilde{\mathbf{x}}} \boldsymbol{\varphi}^\epsilon(s - \alpha), \quad (23)$$

where

$$\frac{\partial \bar{\mathbf{f}}^\epsilon(s, \eta, \boldsymbol{\sigma})}{\partial \mathbf{x}} = \frac{\partial \mathbf{f}(\mathbf{x}(s) + \eta\boldsymbol{\varphi}^\epsilon(s), \mathbf{x}(s - \alpha) + \eta\boldsymbol{\varphi}^\epsilon(s - \alpha), \mathbf{u}(s), \boldsymbol{\sigma})}{\partial \mathbf{x}}, \quad (24)$$

$$\frac{\partial \bar{\mathbf{f}}^\epsilon(s, \eta, \boldsymbol{\sigma})}{\partial \tilde{\mathbf{x}}} = \frac{\partial \mathbf{f}(\mathbf{x}(s) + \eta\boldsymbol{\varphi}^\epsilon(s), \mathbf{x}(s - \alpha) + \eta\boldsymbol{\varphi}^\epsilon(s - \alpha), \mathbf{u}(s), \boldsymbol{\sigma})}{\partial \tilde{\mathbf{x}}}. \quad (25)$$

We can rewrite (23) as follows:

$$\frac{\partial \bar{\mathbf{f}}^\epsilon(s, \eta, \boldsymbol{\sigma})}{\partial \eta} = \Delta_1(s, \eta, \boldsymbol{\sigma}) + \Delta_2(s, \eta, \boldsymbol{\sigma}) + \frac{\partial \bar{\mathbf{f}}^\epsilon(s, 0, \boldsymbol{\sigma})}{\partial \mathbf{x}} \boldsymbol{\varphi}^\epsilon(s) + \frac{\partial \bar{\mathbf{f}}^\epsilon(s, 0, \boldsymbol{\sigma})}{\partial \tilde{\mathbf{x}}} \boldsymbol{\varphi}^\epsilon(s - \alpha), \quad (26)$$

where

$$\Delta_1(s, \eta, \boldsymbol{\sigma}) = \left\{ \frac{\partial \bar{\mathbf{f}}^\epsilon(s, \eta, \boldsymbol{\sigma})}{\partial \mathbf{x}} - \frac{\partial \bar{\mathbf{f}}^\epsilon(s, 0, \boldsymbol{\sigma})}{\partial \mathbf{x}} \right\} \boldsymbol{\varphi}^\epsilon(s), \quad (27)$$

$$\Delta_2(s, \eta, \boldsymbol{\sigma}) = \left\{ \frac{\partial \bar{\mathbf{f}}^\epsilon(s, \eta, \boldsymbol{\sigma})}{\partial \tilde{\mathbf{x}}} - \frac{\partial \bar{\mathbf{f}}^\epsilon(s, 0, \boldsymbol{\sigma})}{\partial \tilde{\mathbf{x}}} \right\} \boldsymbol{\varphi}^\epsilon(s - \alpha). \quad (28)$$

Since \mathbf{f} is continuously differentiable and \mathbf{x} and \mathbf{u} are bounded, the following result is easily established.

Lemma 2. *For each $\boldsymbol{\sigma} \in \mathbb{R}^r$, there exists a corresponding $L_4 > 0$ such that*

$$\left| \frac{\partial \bar{\mathbf{f}}^\epsilon(s, 0, \boldsymbol{\sigma})}{\partial \mathbf{x}} \right| \leq L_4, \quad \left| \frac{\partial \bar{\mathbf{f}}^\epsilon(s, 0, \boldsymbol{\sigma})}{\partial \tilde{\mathbf{x}}} \right| \leq L_4, \quad s \in [0, T], \quad (29)$$

where $|\cdot|$ denotes the natural matrix norm on $\mathbb{R}^{n \times n}$.

We now show that the functions Δ_1 (defined by (27)) and Δ_2 (defined by (28)) are of order ϵ .

Lemma 3. *Let $\delta > 0$ and $\boldsymbol{\sigma} \in \mathbb{R}^r$ be arbitrary. Then for all $\epsilon \in \mathcal{S}$ of sufficiently small magnitude,*

$$|\Delta_1(s, \eta, \boldsymbol{\sigma})| \leq L_3 \delta |\epsilon|, \quad |\Delta_2(s, \eta, \boldsymbol{\sigma})| \leq L_3 \delta |\epsilon|, \quad (s, \eta) \in [0, T] \times [0, 1],$$

where $L_3 > 0$ is as defined in Lemma 1.

Proof. By (21), $\mathbf{x}(s) + \eta \boldsymbol{\varphi}^\epsilon(s) \rightarrow \mathbf{x}(s)$ and $\mathbf{x}(s - \alpha) + \eta \boldsymbol{\varphi}^\epsilon(s - \alpha) \rightarrow \mathbf{x}(s - \alpha)$ uniformly on $[0, T]$ as $\epsilon \rightarrow 0$. Hence, since \mathbf{f} is continuously differentiable and \mathbf{x}^ϵ is uniformly bounded with respect to ϵ , there exists an $\epsilon' > 0$ such that for all $\epsilon \in \mathcal{S}$ satisfying $|\epsilon| < \epsilon'$,

$$\begin{aligned} \left| \frac{\partial \bar{\mathbf{f}}^\epsilon(s, \eta, \boldsymbol{\sigma})}{\partial \mathbf{x}} - \frac{\partial \bar{\mathbf{f}}^\epsilon(s, 0, \boldsymbol{\sigma})}{\partial \mathbf{x}} \right| &< \delta, \quad (s, \eta) \in [0, T] \times [0, 1], \\ \left| \frac{\partial \bar{\mathbf{f}}^\epsilon(s, \eta, \boldsymbol{\sigma})}{\partial \tilde{\mathbf{x}}} - \frac{\partial \bar{\mathbf{f}}^\epsilon(s, 0, \boldsymbol{\sigma})}{\partial \tilde{\mathbf{x}}} \right| &< \delta, \quad (s, \eta) \in [0, T] \times [0, 1]. \end{aligned}$$

By taking the norm of (27)-(28), and then using the above inequalities together with (21), we obtain the desired result. \square

Let a and b be given constants such that $a, b \in [0, T]$. Define

$$\rho_\epsilon(a, b, \boldsymbol{\sigma}) = \int_a^b \{ \bar{\mathbf{f}}^\epsilon(s, 1, \boldsymbol{\sigma}) - \bar{\mathbf{f}}^\epsilon(s, 0, \boldsymbol{\sigma}) \} ds. \quad (30)$$

Our final preliminary result is stated and proved below.

Lemma 4. *For each $\boldsymbol{\sigma} \in \mathbb{R}^r$, there exists a corresponding $L_5 > 0$ such that*

$$|\rho_\epsilon(a, b, \boldsymbol{\sigma})| \leq L_5 |b - a| \cdot |\epsilon| + L_5 (b - a)^2 + L_5 \int_{\min\{a, b\}}^{\max\{a, b\}} |\mathbf{u}(s) - \mathbf{u}(a)| ds.$$

Proof. From (30),

$$\rho_\epsilon(a, b, \boldsymbol{\sigma}) = \int_{\min\{a, b\}}^{\max\{a, b\}} \{ \bar{\mathbf{f}}^\epsilon(s, 1, \boldsymbol{\sigma}) - \bar{\mathbf{f}}^\epsilon(s, 0, \boldsymbol{\sigma}) \} ds + \int_{\min\{a, b\}}^{\max\{a, b\}} \{ \bar{\mathbf{f}}^\epsilon(s, 0, \boldsymbol{\sigma}) - \bar{\mathbf{f}}^\epsilon(a, 0, \boldsymbol{\sigma}) \} ds.$$

Thus,

$$\begin{aligned} |\rho_\epsilon(a, b, \boldsymbol{\sigma})| &\leq \int_{\min\{a, b\}}^{\max\{a, b\}} |\bar{\mathbf{f}}^\epsilon(s, 1, \boldsymbol{\sigma}) - \bar{\mathbf{f}}^\epsilon(s, 0, \boldsymbol{\sigma})| ds \\ &\quad + \int_{\min\{a, b\}}^{\max\{a, b\}} |\bar{\mathbf{f}}^\epsilon(s, 0, \boldsymbol{\sigma}) - \bar{\mathbf{f}}^\epsilon(a, 0, \boldsymbol{\sigma})| ds. \end{aligned} \quad (31)$$

Consider the first integrand on the right-hand side of (31). Using (26) and (29) yields

$$\begin{aligned} |\bar{\mathbf{f}}^\epsilon(s, 1, \boldsymbol{\sigma}) - \bar{\mathbf{f}}^\epsilon(s, 0, \boldsymbol{\sigma})| &\leq \int_0^1 \left| \frac{\partial \bar{\mathbf{f}}^\epsilon(s, \eta, \boldsymbol{\sigma})}{\partial \eta} \right| d\eta \\ &\leq \int_0^1 \{ |\Delta_1(s, \eta, \boldsymbol{\sigma})| + |\Delta_2(s, \eta, \boldsymbol{\sigma})| \} d\eta \\ &\quad + L_4 |\boldsymbol{\varphi}^\epsilon(s)| + L_4 |\boldsymbol{\varphi}^\epsilon(s - \alpha)|. \end{aligned}$$

By Lemma 1 and Lemma 3 with $\delta = 1$, we see that for all $\epsilon \in \mathcal{S}$ of sufficiently small magnitude,

$$|\bar{\mathbf{f}}^\epsilon(s, 1, \boldsymbol{\sigma}) - \bar{\mathbf{f}}^\epsilon(s, 0, \boldsymbol{\sigma})| \leq 2L_3|\epsilon| + 2L_3L_4|\epsilon|. \quad (32)$$

Now, consider the second integrand on the right-hand side of (31). Since \mathbf{f} is continuously differentiable and \mathbf{x}^ϵ is uniformly bounded with respect to ϵ (recall (20)), there exists a constant $M_3 > 0$ such that

$$|\bar{\mathbf{f}}^\epsilon(s, 0, \boldsymbol{\sigma}) - \bar{\mathbf{f}}^\epsilon(a, 0, \boldsymbol{\sigma})| \leq M_3|\mathbf{x}(s) - \mathbf{x}(a)| + M_3|\mathbf{x}(s - \alpha) - \mathbf{x}(a - \alpha)| + M_3|\mathbf{u}(s) - \mathbf{u}(a)|.$$

Recall that $\dot{\mathbf{x}}(s) = \boldsymbol{\psi}(s)$ for almost all $s \in (-\infty, T]$, where $\boldsymbol{\psi}$ is as defined in Subsection 3.1. Thus,

$$\begin{aligned} |\bar{\mathbf{f}}^\epsilon(s, 0, \boldsymbol{\sigma}) - \bar{\mathbf{f}}^\epsilon(a, 0, \boldsymbol{\sigma})| &\leq M_3 \int_{\min\{a,s\}}^{\max\{a,s\}} |\boldsymbol{\psi}(\eta)| d\eta + M_3 \int_{\min\{a,s\}-\alpha}^{\max\{a,s\}-\alpha} |\boldsymbol{\psi}(\eta)| d\eta + M_3|\mathbf{u}(s) - \mathbf{u}(a)| \\ &\leq M_3M_4|s - a| + M_3M_4|s - a| + M_3|\mathbf{u}(s) - \mathbf{u}(a)|, \end{aligned} \quad (33)$$

where $M_4 = \max_{\eta \in [-\alpha_{\max}, T]} |\boldsymbol{\psi}(\eta)|$. Substituting (32) and (33) into (31) gives

$$|\rho_\epsilon(a, b, \boldsymbol{\sigma})| \leq (2L_3 + 2L_3L_4)|b - a| \cdot |\epsilon| + 2M_3M_4(b - a)^2 + M_3 \int_{\min\{a,b\}}^{\max\{a,b\}} |\mathbf{u}(s) - \mathbf{u}(a)| ds.$$

Taking $L_5 = \max\{2L_3 + 2L_3L_4, M_3, 2M_3M_4\}$ completes the proof. \square

3.2.2 Main result

Armed with Lemmas 1-4, we are now ready to derive the state variation with respect to the input-delay β . Define

$$\mathcal{I} = \{t_i + \beta, i = 0, \dots, p\}.$$

First, consider the following auxiliary system:

$$\begin{aligned} \dot{\boldsymbol{\Gamma}}(t) &= \frac{\partial \mathbf{f}(\mathbf{x}(t), \mathbf{x}(t - \alpha), \mathbf{u}(t), \mathbf{u}(t - \beta))}{\partial \mathbf{x}} \boldsymbol{\Gamma}(t) \\ &\quad + \frac{\partial \mathbf{f}(\mathbf{x}(t), \mathbf{x}(t - \alpha), \mathbf{u}(t), \mathbf{u}(t - \beta))}{\partial \tilde{\mathbf{x}}} \boldsymbol{\Gamma}(t - \alpha), \end{aligned} \quad (34)$$

where, for each $t \in \mathcal{I} \cap (0, T]$,

$$\lim_{t \rightarrow (t_i + \beta)^+} \boldsymbol{\Gamma}(t) = \lim_{t \rightarrow (t_i + \beta)^-} \boldsymbol{\Gamma}(t) + \bar{\mathbf{f}}^\epsilon(t_i + \beta, 0, \boldsymbol{\sigma}^i) - \bar{\mathbf{f}}^\epsilon(t_i + \beta, 0, \boldsymbol{\sigma}^{i+1}), \quad (35)$$

and

$$\boldsymbol{\Gamma}(t) = \mathbf{0}, \quad t \leq 0. \quad (36)$$

Let $\boldsymbol{\Gamma}(\cdot | \alpha, \beta)$ denote the unique right continuous solution of (34)-(36). We have the following important result.

Theorem 2. Let $(\alpha, \beta) \in [\alpha_{\min}, \alpha_{\max}] \times [\beta_{\min}, \beta_{\max})$ be a fixed pair such that

$$t_i + \beta \notin \{0\} \cup \{t_j, j = 0, \dots, p\}, \quad i = 0, \dots, p.$$

Furthermore, consider a fixed time point $t \in (t_{i-1} + \beta, t_i + \beta) \cap (0, T]$, where $i \in \{1, \dots, p\}$.

Then

$$\lim_{\epsilon \rightarrow 0^+} \epsilon^{-1} \varphi^\epsilon(t) = \Gamma(t|\alpha, \beta), \quad (37)$$

where φ^ϵ is as defined in (18).

Proof. Let

$$a_i = \max\{t_{i-1} + \beta, 0\}.$$

Then

$$\mathbf{x}(t) = \mathbf{x}(a_i) + \int_{a_i}^t \bar{\mathbf{f}}^\epsilon(s, 0, \boldsymbol{\sigma}^i) ds. \quad (38)$$

Let $\epsilon \in \mathcal{S}$ be sufficiently small so that $0 < \epsilon < \min\{t_j - t_{j-1}\}_{j=1}^p$ and $t > t_{i-1} + \beta + \epsilon$. Define

$$a_i^\epsilon = \max\{t_{i-1} + \beta + \epsilon, 0\}.$$

Then

$$\mathbf{x}^\epsilon(t) = \mathbf{x}^\epsilon(a_i^\epsilon) + \int_{a_i^\epsilon}^t \bar{\mathbf{f}}^\epsilon(s, 1, \boldsymbol{\sigma}^i) ds. \quad (39)$$

We can write (39) as follows:

$$\begin{aligned} \mathbf{x}^\epsilon(t) &= \mathbf{x}^\epsilon(a_i) + \int_{a_i}^{a_i^\epsilon} \bar{\mathbf{f}}^\epsilon(s, 1, \boldsymbol{\sigma}^{i-1}) ds + \int_{a_i^\epsilon}^t \bar{\mathbf{f}}^\epsilon(s, 1, \boldsymbol{\sigma}^i) ds \\ &= \mathbf{x}^\epsilon(a_i) + \int_{a_i}^{a_i^\epsilon} \{\bar{\mathbf{f}}^\epsilon(s, 1, \boldsymbol{\sigma}^{i-1}) - \bar{\mathbf{f}}^\epsilon(s, 1, \boldsymbol{\sigma}^i)\} ds + \int_{a_i}^t \bar{\mathbf{f}}^\epsilon(s, 1, \boldsymbol{\sigma}^i) ds, \end{aligned} \quad (40)$$

where $\boldsymbol{\sigma}^{i-1}$ is arbitrary if $i = 1$ (in this case, we must have $a_i^\epsilon = a_i = 0$ when ϵ is sufficiently small, because $\beta < \beta_{\max}$). From (38) and (40), we have

$$\begin{aligned} \varphi^\epsilon(t) &= \mathbf{x}^\epsilon(t) - \mathbf{x}(t) \\ &= \varphi^\epsilon(a_i) + \int_{a_i}^t \{\bar{\mathbf{f}}^\epsilon(s, 1, \boldsymbol{\sigma}^i) - \bar{\mathbf{f}}^\epsilon(s, 0, \boldsymbol{\sigma}^i)\} ds + \int_{a_i}^{a_i^\epsilon} \{\bar{\mathbf{f}}^\epsilon(s, 1, \boldsymbol{\sigma}^{i-1}) - \bar{\mathbf{f}}^\epsilon(s, 1, \boldsymbol{\sigma}^i)\} ds. \end{aligned}$$

Thus,

$$\begin{aligned} \varphi^\epsilon(t) &= \varphi^\epsilon(a_i) + \int_{a_i}^t \{\bar{\mathbf{f}}^\epsilon(s, 1, \boldsymbol{\sigma}^i) - \bar{\mathbf{f}}^\epsilon(s, 0, \boldsymbol{\sigma}^i)\} ds - \rho_\epsilon(a_i, a_i^\epsilon, \boldsymbol{\sigma}^i) + \rho_\epsilon(a_i, a_i^\epsilon, \boldsymbol{\sigma}^{i-1}) \\ &\quad + (a_i^\epsilon - a_i) \{\bar{\mathbf{f}}^\epsilon(a_i, 0, \boldsymbol{\sigma}^{i-1}) - \bar{\mathbf{f}}^\epsilon(a_i, 0, \boldsymbol{\sigma}^i)\}, \end{aligned}$$

where ρ_ϵ is as defined in (30). By the fundamental theorem of calculus,

$$\begin{aligned} \varphi^\epsilon(t) &= \varphi^\epsilon(a_i) + \int_{a_i}^t \int_0^1 \frac{\partial \bar{\mathbf{f}}^\epsilon(s, \eta, \boldsymbol{\sigma}^i)}{\partial \eta} d\eta ds - \rho_\epsilon(a_i, a_i^\epsilon, \boldsymbol{\sigma}^i) + \rho_\epsilon(a_i, a_i^\epsilon, \boldsymbol{\sigma}^{i-1}) \\ &\quad + (a_i^\epsilon - a_i) \{ \bar{\mathbf{f}}^\epsilon(a_i, 0, \boldsymbol{\sigma}^{i-1}) - \bar{\mathbf{f}}^\epsilon(a_i, 0, \boldsymbol{\sigma}^i) \}. \end{aligned}$$

Using (26),

$$\begin{aligned} \varphi^\epsilon(t) &= \varphi^\epsilon(a_i) + \int_{a_i}^t \int_0^1 \{ \Delta_1(s, \eta, \boldsymbol{\sigma}^i) + \Delta_2(s, \eta, \boldsymbol{\sigma}^i) \} d\eta ds \\ &\quad + \int_{a_i}^t \frac{\partial \bar{\mathbf{f}}^\epsilon(s, 0, \boldsymbol{\sigma}^i)}{\partial \mathbf{x}} \varphi^\epsilon(s) ds + \int_{a_i}^t \frac{\partial \bar{\mathbf{f}}^\epsilon(s, 0, \boldsymbol{\sigma}^i)}{\partial \tilde{\mathbf{x}}} \varphi^\epsilon(s - \alpha) ds - \rho_\epsilon(a_i, a_i^\epsilon, \boldsymbol{\sigma}^i) \\ &\quad + \rho_\epsilon(a_i, a_i^\epsilon, \boldsymbol{\sigma}^{i-1}) + (a_i^\epsilon - a_i) \{ \bar{\mathbf{f}}^\epsilon(a_i, 0, \boldsymbol{\sigma}^{i-1}) - \bar{\mathbf{f}}^\epsilon(a_i, 0, \boldsymbol{\sigma}^i) \}. \end{aligned} \quad (41)$$

We can express the solution of the auxiliary system as follows:

$$\Gamma(t) = \Gamma(a_i^+) + \int_{a_i}^t \frac{\partial \bar{\mathbf{f}}^\epsilon(s, 0, \boldsymbol{\sigma}^i)}{\partial \mathbf{x}} \Gamma(s) ds + \int_{a_i}^t \frac{\partial \bar{\mathbf{f}}^\epsilon(s, 0, \boldsymbol{\sigma}^i)}{\partial \tilde{\mathbf{x}}} \Gamma(s - \alpha) ds. \quad (42)$$

Thus, from Lemma 2 and equations (41) and (42),

$$\begin{aligned} |\epsilon^{-1} \varphi^\epsilon(t) - \Gamma(t)| &\leq |\gamma_i(\epsilon)| + \epsilon^{-1} \int_{a_i}^t \int_0^1 \{ |\Delta_1(s, \eta, \boldsymbol{\sigma}^i)| + |\Delta_2(s, \eta, \boldsymbol{\sigma}^i)| \} d\eta ds \\ &\quad + \int_{a_i}^t L_4 |\epsilon^{-1} \varphi^\epsilon(s) - \Gamma(s)| ds + \int_{a_i}^t L_4 |\epsilon^{-1} \varphi^\epsilon(s - \alpha) - \Gamma(s - \alpha)| ds \\ &\quad + \epsilon^{-1} |\rho_\epsilon(a_i, a_i^\epsilon, \boldsymbol{\sigma}^i)| + \epsilon^{-1} |\rho_\epsilon(a_i, a_i^\epsilon, \boldsymbol{\sigma}^{i-1})|, \end{aligned}$$

where L_4 is the constant defined Lemma 2 and

$$\gamma_i(\epsilon) = \epsilon^{-1} \varphi^\epsilon(a_i) - \Gamma(a_i^+) + \epsilon^{-1} (a_i^\epsilon - a_i) \{ \bar{\mathbf{f}}^\epsilon(a_i, 0, \boldsymbol{\sigma}^{i-1}) - \bar{\mathbf{f}}^\epsilon(a_i, 0, \boldsymbol{\sigma}^i) \}. \quad (43)$$

Recall that $a_i^\epsilon - a_i \leq \epsilon$ and $t_{i-1} + \beta \neq t_j$ for all j . Thus, we may assume that ϵ is sufficiently small so that $\mathbf{u}(s) = \mathbf{u}(a_i)$ for all $s \in [a_i, a_i^\epsilon]$. It then follows from Lemma 4 that

$$|\rho_\epsilon(a_i, a_i^\epsilon, \boldsymbol{\sigma}^{i-1})| \leq 2L_5' \epsilon^2, \quad |\rho_\epsilon(a_i, a_i^\epsilon, \boldsymbol{\sigma}^i)| \leq 2L_5'' \epsilon^2,$$

where L_5' and L_5'' are the constants in Lemma 4 corresponding to $\boldsymbol{\sigma}^{i-1}$ and $\boldsymbol{\sigma}^i$, respectively. By the above inequalities and Lemma 3, assuming that ϵ is sufficiently small,

$$\begin{aligned} |\epsilon^{-1} \varphi^\epsilon(t) - \Gamma(t)| &\leq 2TL_3\delta + 2L_5'\epsilon + 2L_5''\epsilon + |\gamma_i(\epsilon)| + \int_{a_i}^t L_4 |\epsilon^{-1} \varphi^\epsilon(s) - \Gamma(s)| ds \\ &\quad + \int_{a_i}^t L_4 |\epsilon^{-1} \varphi^\epsilon(s - \alpha) - \Gamma(s - \alpha)| ds, \end{aligned} \quad (44)$$

where $\delta > 0$ is arbitrary and L_3 is the constant defined in Lemma 1. Performing a change of variable in the second integral in the right-hand side of (44) yields

$$\begin{aligned} |\epsilon^{-1} \varphi^\epsilon(t) - \Gamma(t)| &\leq 2TL_3\delta + 4L_5\epsilon + |\gamma_i(\epsilon)| + \int_{a_i}^t L_4 |\epsilon^{-1} \varphi^\epsilon(s) - \Gamma(s)| ds \\ &\quad + \int_{a_i - \alpha}^{t - \alpha} L_4 |\epsilon^{-1} \varphi^\epsilon(s) - \Gamma(s)| ds \\ &\leq 2TL_3\delta + 4L_5\epsilon + |\gamma_i(\epsilon)| + \mu_i(\epsilon) + \int_{a_i}^t 2L_4 |\epsilon^{-1} \varphi^\epsilon(s) - \Gamma(s)| ds, \end{aligned}$$

where $L_5 = \max\{L'_5, L''_5\}$ and

$$\mu_i(\epsilon) = \int_{a_i - \alpha}^{a_i} L_4 |\epsilon^{-1} \varphi^\epsilon(s) - \Gamma(s)| ds.$$

Assuming that δ is sufficiently small so that $a_i + \delta \leq t$,

$$\begin{aligned} |\epsilon^{-1} \varphi^\epsilon(t) - \Gamma(t)| &\leq 2TL_3\delta + 4L_5\epsilon + |\gamma_i(\epsilon)| + \mu_i(\epsilon) + \int_{a_i}^{\delta} 2L_4 |\epsilon^{-1} \varphi^\epsilon(s) - \Gamma(s)| ds \\ &+ \int_{\delta}^t 2L_4 |\epsilon^{-1} \varphi^\epsilon(s) - \Gamma(s)| ds, \end{aligned} \quad (45)$$

Now since Γ is a piecewise continuous function, there exists a constant $M_1 > 0$ such that

$$|\Gamma(s)| \leq M_1, \quad s \in (-\infty, T].$$

Therefore, it follows from Lemma 1 that for all sufficiently small $\epsilon > 0$,

$$|\epsilon^{-1} \varphi^\epsilon(s) - \Gamma(s)| \leq L_3 + M_1, \quad s \in (-\infty, T]. \quad (46)$$

Substituting (46) into (45) gives

$$\begin{aligned} |\epsilon^{-1} \varphi^\epsilon(t) - \Gamma(t)| &\leq 2TL_3\delta + 4L_5\epsilon + |\gamma_i(\epsilon)| + \mu_i(\epsilon) + 2L_4(L_3 + M_1)\delta \\ &+ \int_{\delta}^t 2L_4 |\epsilon^{-1} \varphi^\epsilon(s) - \Gamma(s)| ds, \end{aligned} \quad (47)$$

Note that this inequality holds for all $t \in [a_i + \delta, t_i + \beta)$ and $t = (t_i + \beta)^-$, uniformly with respect to $\epsilon \leq \delta$. Thus, by the Gronwall-Bellman Lemma [17],

$$|\epsilon^{-1} \varphi^\epsilon(t) - \Gamma(t)| \leq (2TL_3\delta + 4L_5\epsilon + |\gamma_i(\epsilon)| + \mu_i(\epsilon) + 2L_4(L_3 + M_1)\delta) \exp\{2L_4T\}. \quad (48)$$

This inequality holds for all ϵ of sufficiently small magnitude.

Now, suppose that $t \in (t_{i-1} + \beta, t_i + \beta) \cap (0, T]$ for $i = \min\{j : t_j + \beta > 0\}$. Then $a_i = 0$, and thus by (2) and (36),

$$\mu_i(\epsilon) = \int_{-\alpha}^0 L_4 |\epsilon^{-1} \varphi^\epsilon(s) - \Gamma(s)| ds = 0.$$

Since by assumption $t_{i-1} + \beta < 0$, $a_i^\epsilon = a_i = 0$ for all sufficiently small ϵ . Thus,

$$\gamma_i(\epsilon) = \epsilon^{-1} \varphi^\epsilon(0) - \Gamma(0^+) = \mathbf{0}.$$

Substituting $\mu_i(\epsilon) = 0$ and $\gamma_i(\epsilon) = \mathbf{0}$ into (48) gives

$$|\epsilon^{-1} \varphi^\epsilon(t) - \Gamma(t)| \leq (2TL_3\delta + 4L_5\epsilon + 2L_4(L_3 + M_1)\delta) \exp\{2L_4T\}. \quad (49)$$

Since $\delta > 0$ was chosen arbitrarily and ϵ can be made arbitrarily small, this shows that (37) holds for $i = \min\{j : t_j + \beta > 0\}$. Moreover, the derivation leading to (49) shows that (37) also holds for $t = (t_i + \beta)^-$. It is also clear that (37) holds for all $t \in (-\infty, 0]$.

Now, suppose that (37) holds for all $t \in (-\infty, t_k + \beta) \setminus \{t_j + \beta\}_{j=0}^k$ and $t = (t_k + \beta)^-$, where

$$\min\{j : t_j + \beta > 0\} \leq k \leq p - 1. \quad (50)$$

We will show that (37) holds for all $t \in (t_k + \beta, t_{k+1} + \beta)$ and $t = (t_{k+1} + \beta)^-$. The result will then follow by induction.

Let $t \in (t_k + \beta, t_{k+1} + \beta)$, where k satisfies (50). By our inductive hypothesis, for almost all $s \in (-\infty, t_k + \beta)$,

$$\lim_{\epsilon \rightarrow 0^+} \epsilon^{-1} \varphi^\epsilon(s) = \Gamma(s). \quad (51)$$

In view of (46) and (51), we can apply Lebesgue's dominated convergence theorem to obtain

$$\lim_{\epsilon \rightarrow 0^+} \mu_{k+1}(\epsilon) = \lim_{\epsilon \rightarrow 0^+} \int_{t_k + \beta - \alpha}^{t_k + \beta} L_4 |\epsilon^{-1} \varphi^\epsilon(s) - \Gamma(s)| ds = 0. \quad (52)$$

Furthermore,

$$\begin{aligned} \gamma_{k+1}(\epsilon) &= \epsilon^{-1} \varphi^\epsilon(t_k + \beta) - \Gamma((t_k + \beta)^+) + \bar{\mathbf{f}}^\epsilon(t_k + \beta, 0, \boldsymbol{\sigma}^k) - \bar{\mathbf{f}}^\epsilon(t_k + \beta, 0, \boldsymbol{\sigma}^{k+1}) \\ &= \epsilon^{-1} \varphi^\epsilon(t_k + \beta) - \Gamma((t_k + \beta)^-). \end{aligned}$$

Thus, by our inductive hypothesis,

$$\lim_{\epsilon \rightarrow 0^+} \gamma_{k+1}(\epsilon) = \mathbf{0}. \quad (53)$$

By combining equations (52) and (53) with (48) for $i = k + 1$, we see that (37) holds for $t \in (t_k + \beta, t_{k+1} + \beta)$. Similar arguments show that (37) also holds for $t = (t_{k+1} + \beta)^-$. The proof then follows by induction. \square

Theorem 2 shows that $\epsilon^{-1} \varphi^\epsilon \rightarrow \Gamma(\cdot | \alpha, \beta)$ as $\epsilon \rightarrow 0^+$. We now derive the analogous result for $\epsilon \rightarrow 0^-$.

Theorem 3. *Let $(\alpha, \beta) \in [\alpha_{\min}, \alpha_{\max}] \times (\beta_{\min}, \beta_{\max}]$ be a fixed pair such that*

$$t_i + \beta \notin \{0\} \cup \{t_j, j = 0, \dots, p\}, \quad i = 0, \dots, p.$$

Furthermore, consider a fixed time point $t \in (t_{i-1} + \beta, t_i + \beta) \cap (0, T]$, where $i \in \{1, \dots, p\}$. Then

$$\lim_{\epsilon \rightarrow 0^-} \epsilon^{-1} \varphi^\epsilon(t) = \Gamma(t | \alpha, \beta). \quad (54)$$

Proof. Let a_i and a_i^ϵ be as defined in the proof of Theorem 2. Furthermore, let $\epsilon \in \mathcal{S}$ such that $\min\{t_{j-1} - t_j\}_{j=1}^p < \epsilon < 0$ and $t < t_i + \beta + \epsilon$. Then

$$\mathbf{x}(t) = \mathbf{x}(a_i^\epsilon) + \int_{a_i^\epsilon}^{a_i} \bar{\mathbf{f}}^\epsilon(s, 0, \boldsymbol{\sigma}^{i-1}) ds + \int_{a_i}^t \bar{\mathbf{f}}^\epsilon(s, 0, \boldsymbol{\sigma}^i) ds,$$

where $\boldsymbol{\sigma}^{i-1}$ is arbitrary if $i = 1$ (in this case, we must have $a_i = a_i^\epsilon = 0$). Moreover,

$$\mathbf{x}^\epsilon(t) = \mathbf{x}^\epsilon(a_i^\epsilon) + \int_{a_i^\epsilon}^{a_i} \bar{\mathbf{f}}^\epsilon(s, 1, \boldsymbol{\sigma}^i) ds + \int_{a_i}^t \bar{\mathbf{f}}^\epsilon(s, 1, \boldsymbol{\sigma}^i) ds.$$

Thus,

$$\begin{aligned}\varphi^\epsilon(t) &= \mathbf{x}^\epsilon(t) - \mathbf{x}(t) \\ &= \varphi^\epsilon(a_i^\epsilon) + \int_{a_i^\epsilon}^{a_i} (\bar{\mathbf{f}}^\epsilon(s, 1, \boldsymbol{\sigma}^i) - \bar{\mathbf{f}}^\epsilon(s, 0, \boldsymbol{\sigma}^{i-1})) ds + \int_{a_i}^t (\bar{\mathbf{f}}^\epsilon(s, 1, \boldsymbol{\sigma}^i) - \bar{\mathbf{f}}^\epsilon(s, 0, \boldsymbol{\sigma}^i)) ds.\end{aligned}$$

This equation can be rewritten as follows:

$$\begin{aligned}\varphi^\epsilon(t) &= \varphi^\epsilon(a_i^\epsilon) - \rho_\epsilon(a_i, a_i^\epsilon, \boldsymbol{\sigma}^i) + (a_i - a_i^\epsilon) \{ \bar{\mathbf{f}}^\epsilon(a_i, 0, \boldsymbol{\sigma}^i) - \bar{\mathbf{f}}^\epsilon(a_i, 0, \boldsymbol{\sigma}^{i-1}) \} \\ &\quad + \int_{a_i^\epsilon}^{a_i} \{ \bar{\mathbf{f}}^\epsilon(a_i, 0, \boldsymbol{\sigma}^{i-1}) - \bar{\mathbf{f}}^\epsilon(s, 0, \boldsymbol{\sigma}^{i-1}) \} ds + \int_{a_i}^t \{ \bar{\mathbf{f}}^\epsilon(s, 1, \boldsymbol{\sigma}^i) - \bar{\mathbf{f}}^\epsilon(s, 0, \boldsymbol{\sigma}^i) \} ds,\end{aligned}$$

where ρ_ϵ is as defined in (30). Using the fundamental theorem of calculus and (26),

$$\begin{aligned}\varphi^\epsilon(t) &= \varphi^\epsilon(a_i^\epsilon) - \rho_\epsilon(a_i, a_i^\epsilon, \boldsymbol{\sigma}^i) + \int_{a_i^\epsilon}^{a_i} \{ \bar{\mathbf{f}}^\epsilon(a_i, 0, \boldsymbol{\sigma}^{i-1}) - \bar{\mathbf{f}}^\epsilon(s, 0, \boldsymbol{\sigma}^{i-1}) \} ds \\ &\quad + (a_i - a_i^\epsilon) \{ \bar{\mathbf{f}}^\epsilon(a_i, 0, \boldsymbol{\sigma}^i) - \bar{\mathbf{f}}^\epsilon(a_i, 0, \boldsymbol{\sigma}^{i-1}) \} + \int_{a_i}^t \frac{\partial \bar{\mathbf{f}}^\epsilon(s, 0, \boldsymbol{\sigma}^i)}{\partial \mathbf{x}} \varphi^\epsilon(s) ds \\ &\quad + \int_{a_i}^t \frac{\partial \bar{\mathbf{f}}^\epsilon(s, 0, \boldsymbol{\sigma}^i)}{\partial \tilde{\mathbf{x}}} \varphi^\epsilon(s - \alpha) ds + \int_{a_i}^t \int_0^1 \{ \Delta_1(s, \eta, \boldsymbol{\sigma}^i) + \Delta_2(s, \eta, \boldsymbol{\sigma}^i) \} d\eta ds.\end{aligned}$$

Recall that $t_i + \beta \neq t_j$ for all i and j . Thus, we may assume that ϵ is sufficiently small so that $\mathbf{u}(s) = \mathbf{u}(a_i^\epsilon)$ for all $s \in [a_i^\epsilon, a_i]$. It then follows from Lemma 4 that

$$|\rho_\epsilon(a_i, a_i^\epsilon, \boldsymbol{\sigma}^{i-1})| \leq 2L_5|\epsilon|^2.$$

Furthermore, assuming that ϵ is sufficiently small, by using the similar arguments to those in the proof of Lemma 4, one can show that there exists a constant $M_5 > 0$ such that

$$\int_{a_i^\epsilon}^{a_i} |\bar{\mathbf{f}}^\epsilon(a_i, 0, \boldsymbol{\sigma}^{i-1}) - \bar{\mathbf{f}}^\epsilon(s, 0, \boldsymbol{\sigma}^{i-1})| ds \leq M_5\epsilon^2.$$

Hence, as in the proof of Theorem 2,

$$\begin{aligned}|\epsilon^{-1}\varphi^\epsilon(t) - \mathbf{\Gamma}(t)| &\leq |\gamma_i(\epsilon)| + M_5|\epsilon| + 2L_5|\epsilon| + 2L_3T\delta + \int_{a_i}^t L_4|\epsilon^{-1}\varphi^\epsilon(s) - \mathbf{\Gamma}(s)| ds \\ &\quad + \int_{a_i}^t L_4|\epsilon^{-1}\varphi^\epsilon(s - \alpha) - \mathbf{\Gamma}(s - \alpha)| ds,\end{aligned}\tag{55}$$

where $\delta > 0$ is arbitrary and

$$\gamma_i(\epsilon) = \epsilon^{-1}\varphi^\epsilon(a_i^\epsilon) - \mathbf{\Gamma}(a_i^+) + \epsilon^{-1}(a_i - a_i^\epsilon) \{ \bar{\mathbf{f}}^\epsilon(a_i, 0, \boldsymbol{\sigma}^i) - \bar{\mathbf{f}}^\epsilon(a_i, 0, \boldsymbol{\sigma}^{i-1}) \}.$$

Simplifying (55) gives

$$|\epsilon^{-1}\varphi^\epsilon(t) - \mathbf{\Gamma}(t)| \leq |\gamma_i(\epsilon)| + M_5|\epsilon| + 2L_5|\epsilon| + 2L_3T\delta + \mu_i(\epsilon) + \int_{a_i}^t 2L_4|\epsilon^{-1}\varphi^\epsilon(s) - \mathbf{\Gamma}(s)| ds,$$

where

$$\mu_i(\epsilon) = \int_{a_i - \alpha}^{a_i} L_4|\epsilon^{-1}\varphi^\epsilon(s) - \mathbf{\Gamma}(s)| ds.$$

Assuming that $\delta > 0$ is sufficiently small so that $a_i + \delta \leq t$,

$$\begin{aligned} |\epsilon^{-1}\varphi^\epsilon(t) - \mathbf{\Gamma}(t)| &\leq |\gamma_i(\epsilon)| + M_5|\epsilon| + 2L_5|\epsilon| + 2L_3T\delta + \mu_i(\epsilon) \\ &\quad + \int_{a_i}^{a_i+\delta} 2L_4|\epsilon^{-1}\varphi^\epsilon(s) - \mathbf{\Gamma}(s)|ds + \int_{a_i+\delta}^t 2L_4|\epsilon^{-1}\varphi^\epsilon(s) - \mathbf{\Gamma}(s)|ds, \end{aligned} \quad (56)$$

Now since $\mathbf{\Gamma}$ is a piecewise continuous function, there exists a constant $M_1 > 0$ such that

$$|\mathbf{\Gamma}(s)| \leq M_1, \quad s \in (-\infty, T].$$

Therefore, it follows from Lemma 1 that for all sufficiently small $\epsilon < 0$,

$$|\epsilon^{-1}\varphi^\epsilon(s) - \mathbf{\Gamma}(s)| \leq L_3 + M_1, \quad s \in (-\infty, T]. \quad (57)$$

Substituting (57) into (56) gives

$$\begin{aligned} |\epsilon^{-1}\varphi^\epsilon(t) - \mathbf{\Gamma}(t)| &\leq M_5|\epsilon| + 2L_5|\epsilon| + 2L_3T\delta + |\gamma_i(\epsilon)| + \mu_i(\epsilon) + 2L_4(L_3 + M_1)\delta \\ &\quad + \int_{\delta}^t 2L_4|\epsilon^{-1}\varphi^\epsilon(s) - \mathbf{\Gamma}(s)|ds, \end{aligned} \quad (58)$$

Finally, by applying Gronwall's Lemma [17] yields

$$|\epsilon^{-1}\varphi^\epsilon(t) - \mathbf{\Gamma}(t)| \leq (M_5|\epsilon| + 2L_5|\epsilon| + 2L_3T\delta + |\gamma_i(\epsilon)| + \mu_i(\epsilon) + 2L_4(L_3 + M_1)\delta) \exp\{2L_4T\}. \quad (59)$$

This inequality holds for all ϵ of sufficiently small magnitude.

Now, suppose $t \in (t_{i-1} + \beta, t_i + \beta) \cap (0, T]$ for $i = \min\{j : t_j + \beta > 0\}$. Then $a_i = 0$, and thus by (2) and (36),

$$\mu_i(\epsilon) = \int_{-\alpha}^0 L_4|\epsilon^{-1}\varphi^\epsilon(s) - \mathbf{\Gamma}(s)|ds = 0.$$

Also, since $a_i^\epsilon = a_i = 0$,

$$\gamma_i(\epsilon) = \epsilon^{-1}\varphi^\epsilon(0) - \mathbf{\Gamma}(0^+) = \mathbf{0}.$$

Substituting $\mu_i(\epsilon) = 0$ and $\gamma_i(\epsilon) = \mathbf{0}$ into (59) gives

$$|\epsilon^{-1}\varphi^\epsilon(t) - \mathbf{\Gamma}(t)| \leq (L_6|\epsilon| + 2L_5|\epsilon| + 2L_3T\delta) \exp\{2L_4T\}.$$

Since $\delta > 0$ was chosen arbitrary and ϵ can be made arbitrarily small, this shows that (54) holds for $i = \min\{j : t_j + \beta > 0\}$. It is clear that (54) also holds for all $t \in (-\infty, 0]$, and for $t = a_i^+$.

Now, suppose that (54) holds for all $t \in (-\infty, t_k + \beta) \setminus \{t_j + \beta\}_{j=0}^k$ and $t = (t_{k-1} + \beta)^+$, where

$$\min\{j : t_j + \beta > 0\} \leq k \leq p - 1. \quad (60)$$

We will show that (54) holds for all $t \in (t_k + \beta, t_{k+1} + \beta)$ and $t = (t_k + \beta)^+$. The result will then follow by induction.

Let $t \in (t_k + \beta, t_{k+1} + \beta)$, where k satisfies (60) above. By our inductive hypothesis, for almost all $s \in (-\infty, t_k + \beta)$,

$$\lim_{\epsilon \rightarrow 0^-} \epsilon^{-1} \varphi^\epsilon(s) = \Gamma(s). \quad (61)$$

As in the proof of Theorem 2, we can apply Lebesgue's dominated convergence theorem to obtain

$$\lim_{\epsilon \rightarrow 0^+} \mu_{k+1}(\epsilon) = \int_{t_k + \beta - \alpha}^{t_k + \beta} L_4 |\epsilon^{-1} \varphi^\epsilon(s) - \Gamma(s)| ds = 0. \quad (62)$$

We have

$$\begin{aligned} \gamma_{k+1}(\epsilon) &= \epsilon^{-1} \varphi^\epsilon(t_k + \beta + \epsilon) - \Gamma((t_k + \beta)^+) + \bar{\mathbf{f}}^\epsilon(t_k + \beta, 0, \boldsymbol{\sigma}^k) - \bar{\mathbf{f}}^\epsilon(t_k + \beta, 0, \boldsymbol{\sigma}^{k+1}) \\ &= \epsilon^{-1} \varphi^\epsilon(t_k + \beta + \epsilon) - \Gamma((t_k + \beta)^-). \end{aligned}$$

Hence,

$$\begin{aligned} \gamma_{k+1}(\epsilon) &= \epsilon^{-1} \varphi^\epsilon(a_k) - \Gamma(a_k^+) + \epsilon^{-1} \int_{a_k}^{t_k + \beta + \epsilon} \{ \bar{\mathbf{f}}^\epsilon(s, 1, \boldsymbol{\sigma}^k) - \bar{\mathbf{f}}^\epsilon(s, 0, \boldsymbol{\sigma}^k) \} ds \\ &\quad - \int_{a_k}^{t_k + \beta} \frac{\partial \bar{\mathbf{f}}^\epsilon(s, 0, \boldsymbol{\sigma}^k)}{\partial \mathbf{x}} \Gamma(s) ds - \int_{a_k}^{t_k + \beta} \frac{\partial \bar{\mathbf{f}}^\epsilon(s, 0, \boldsymbol{\sigma}^k)}{\partial \tilde{\mathbf{x}}} \Gamma(s - \alpha) ds \\ &= \epsilon^{-1} \varphi^\epsilon(a_k) - \Gamma(a_k^+) + \epsilon^{-1} \int_{a_k}^{t_k + \beta + \epsilon} \int_0^1 \frac{\partial \bar{\mathbf{f}}^\epsilon(s, \eta, \boldsymbol{\sigma}^k)}{\partial \eta} d\eta ds \\ &\quad - \int_{a_k}^{t_k + \beta} \frac{\partial \bar{\mathbf{f}}^\epsilon(s, 0, \boldsymbol{\sigma}^k)}{\partial \mathbf{x}} \Gamma(s) ds - \int_{a_k}^{t_k + \beta} \frac{\partial \bar{\mathbf{f}}^\epsilon(s, 0, \boldsymbol{\sigma}^k)}{\partial \tilde{\mathbf{x}}} \Gamma(s - \alpha) ds. \end{aligned}$$

Using (26), we obtain

$$\begin{aligned} \gamma_{k+1}(\epsilon) &= \epsilon^{-1} \varphi^\epsilon(a_k) - \Gamma(a_k^+) + \epsilon^{-1} \int_{a_k}^{t_k + \beta + \epsilon} \int_0^1 \{ \Delta_1(s, \eta, \boldsymbol{\sigma}^k) + \Delta_2(s, \eta, \boldsymbol{\sigma}^k) \} d\eta ds \\ &\quad + \int_{a_k}^{t_k + \beta + \epsilon} \frac{\partial \bar{\mathbf{f}}^\epsilon(s, 0, \boldsymbol{\sigma}^k)}{\partial \mathbf{x}} \{ \epsilon^{-1} \varphi^\epsilon(s) - \Gamma(s) \} ds \\ &\quad + \int_{a_k}^{t_k + \beta + \epsilon} \frac{\partial \bar{\mathbf{f}}^\epsilon(s, 0, \boldsymbol{\sigma}^k)}{\partial \tilde{\mathbf{x}}} \{ \epsilon^{-1} \varphi^\epsilon(s - \alpha) - \Gamma(s - \alpha) \} ds \\ &\quad - \int_{t_k + \beta - \epsilon}^{t_k + \beta} \frac{\partial \bar{\mathbf{f}}^\epsilon(s, 0, \boldsymbol{\sigma}^k)}{\partial \mathbf{x}} \Gamma(s) ds - \int_{t_k + \beta - \epsilon}^{t_k + \beta} \frac{\partial \bar{\mathbf{f}}^\epsilon(s, 0, \boldsymbol{\sigma}^k)}{\partial \tilde{\mathbf{x}}} \Gamma(s - \alpha) ds. \end{aligned}$$

Thus, using Lemma 2 and Lemma 3,

$$\begin{aligned} |\gamma_{k+1}(\epsilon)| &= |\epsilon^{-1} \varphi^\epsilon(a_k) - \Gamma(a_k^+)| + 2L_3 T \delta + \int_{a_k}^{t_k + \beta + \epsilon} L_4 |\epsilon^{-1} \varphi^\epsilon(s) - \Gamma(s)| ds \\ &\quad + \int_{a_k}^{t_k + \beta + \epsilon} L_4 |\epsilon^{-1} \varphi^\epsilon(s - \alpha) - \Gamma(s - \alpha)| ds + L_4 M_1 |\epsilon| + L_4 M_1 |\epsilon| \\ &\leq |\epsilon^{-1} \varphi^\epsilon(a_k) - \Gamma(a_k^+)| + 2L_3 T \delta + 2L_4 M_1 |\epsilon| \\ &\quad + \int_{a_k}^{t_k + \beta} L_4 |\epsilon^{-1} \varphi^\epsilon(s) - \Gamma(s)| ds + \int_{a_k}^{t_k + \beta} L_4 |\epsilon^{-1} \varphi^\epsilon(s - \alpha) - \Gamma(s - \alpha)| ds \end{aligned}$$

Using the Lebesgue dominated convergence theorem and the induction hypothesis, the two integrals converge to zero as $\epsilon \rightarrow 0^-$. It follows also from the induction hypothesis that the first term converges to zero as $\epsilon \rightarrow 0^-$. Thus,

$$\lim_{\epsilon \rightarrow 0^-} |\gamma_{k+1}(\epsilon)| = 0. \quad (63)$$

Combining equations (62) and (63) with (59), for $i = k + 1$, we see that (54) holds for all $t \in (t_k + \beta, t_{k+1} + \beta)$. The proof then follows by induction. \square

Together, Theorems 2 and 3 show that the state variation with respect to β is given by $\mathbf{\Gamma}(\cdot|\alpha, \beta)$. This is stated formally in the following theorem.

Theorem 4. *Let $(\alpha, \beta) \in [\alpha_{\min}, \alpha_{\max}] \times (\beta_{\min}, \beta_{\max}]$ be a fixed pair such that*

$$t_i + \beta \notin \{0\} \cup \{t_j, j = 0, \dots, p\}, \quad i = 0, \dots, p.$$

Furthermore, consider a fixed time point $t \in (t_{i-1} + \beta, t_i + \beta) \cap (0, T]$, where $i \in \{1, \dots, p\}$.

Then

$$\frac{\partial \mathbf{x}(t|\alpha, \beta)}{\partial \beta} = \mathbf{\Gamma}(t|\alpha, \beta). \quad (64)$$

4 Computational algorithm

In this section, we will develop a computational algorithm for solving Problem (P). Our approach is to view Problem (P) as a nonlinear programming problem in which α and β are the decision variables to be chosen optimally. On this basis, Problem (P) can, in principle, be solved using standard nonlinear programming algorithms such as the SQP method. Such nonlinear programming algorithms typically rely on the partial derivatives of the cost function to compute search directions leading to profitable areas of the research space. Thus, to solve Problem (P) as a nonlinear programming problem, we need to derive the partial derivatives of J with respect to both α and β .

Using Theorem 1, partial derivatives of J with respect to α is given by

$$\begin{aligned} \frac{\partial J(\alpha, \beta)}{\partial \alpha} &= 2 \sum_{j=1}^m (\mathbf{y}(\tau_j|\alpha, \beta) - \bar{\mathbf{y}}^j)^\top \frac{\partial \mathbf{g}(\tau_j|\alpha, \beta)}{\partial \mathbf{x}} \frac{\partial \mathbf{x}(\tau_j|\alpha, \beta)}{\partial \alpha} \\ &\quad + 2 \sum_{j=1}^m (\mathbf{y}(\tau_j|\alpha, \beta) - \bar{\mathbf{y}}^j)^\top \frac{\partial \mathbf{g}(\tau_j|\alpha, \beta)}{\partial \tilde{\mathbf{x}}} \frac{\partial \mathbf{x}(\tau_j - \alpha|\alpha, \beta)}{\partial \alpha} \\ &= 2 \sum_{j=1}^m (\mathbf{y}(\tau_j|\alpha, \beta) - \bar{\mathbf{y}}^j)^\top \frac{\partial \mathbf{g}(\tau_j|\alpha, \beta)}{\partial \mathbf{x}} \mathbf{\Lambda}(\tau_j|\alpha, \beta) \\ &\quad + 2 \sum_{j=1}^m (\mathbf{y}(\tau_j|\alpha, \beta) - \bar{\mathbf{y}}^j)^\top \frac{\partial \mathbf{g}(\tau_j|\alpha, \beta)}{\partial \tilde{\mathbf{x}}} \mathbf{\Lambda}(\tau_j - \alpha|\alpha, \beta), \quad \bar{\tau}_j \in [0, T], \end{aligned} \quad (65)$$

where

$$\frac{\partial \mathbf{g}(\tau_j|\alpha, \beta)}{\partial \mathbf{x}} = \frac{\partial \mathbf{g}(\mathbf{x}(t|\alpha, \beta), \mathbf{x}(t - \alpha|\alpha, \beta))}{\partial \mathbf{x}(t|\alpha, \beta)}, \quad \frac{\partial \mathbf{g}(\tau_j|\alpha, \beta)}{\partial \tilde{\mathbf{x}}} = \frac{\partial \mathbf{g}(\mathbf{x}(t|\alpha, \beta), \mathbf{x}(t - \alpha|\alpha, \beta))}{\partial \mathbf{x}(\tau_j - \alpha|\alpha, \beta)}.$$

Unfortunately, the derivation given above is not valid for the partial derivative of J with respect to β because the state variation with respect to β may not exist in two situations: i) for some values of β (recall that the state variation only exists when $t_i + \beta \notin \{0\} \cup \{t_j, j = 0, \dots, p\}$ for each $i = 0, \dots, p$); ii) at some time points when $t \in \{t_i + \beta\}_{i=0}^p$. To overcome the first difficulty, we first calculate the value of the cost function with the pair (α, β) , if the cost function's value is satisfied, then (α, β) is the optimal solution; or else we perturb β slightly. The perturbation procedure is given by

$$\bar{\beta} = \begin{cases} \beta, & \text{if } t_i + \beta \notin \{0\} \cup \{t_j, j = 0, \dots, p\}, \quad i = 1, \dots, p, \\ \beta + \epsilon, & \text{if } t_i + \beta \in \{0\} \cup \{t_j, j = 0, \dots, p\}, \quad i = 1, \dots, p, \end{cases}$$

where ϵ is a small number such that $\beta + \epsilon \in \mathcal{B}$ and $t_i + \beta \notin \{0\} \cup \{t_j, j = 0, \dots, p\}$ for all $i = 1, \dots, p$. On the other hand, to overcome the second difficulty, we need to consider a modified cost function in which the experimental data are slightly perturbed. The perturbation procedure is designed to ensure that none of the new sample times coincide with points in $\{t_i + \bar{\beta}\}_{i=0}^p \cap (0, T]$. Details are given below.

After arriving at a new delay pair $(\alpha, \bar{\beta})$ at some points during the optimization process, we define the perturbed sample times as follows:

$$\bar{\tau}_j = \begin{cases} \tau_j, & \text{if } \tau_j \notin \{t_i + \bar{\beta}\}_{i=0}^p, \\ \tau_j + \epsilon_j, & \text{if } \tau_j \in \{t_i + \bar{\beta}\}_{i=0}^p, \end{cases} \quad (66)$$

where $\epsilon_j, j = 1, \dots, m$, are small number chosen such that $\tau_j + \epsilon_j \in [0, T] \setminus \{t_i + \bar{\beta}\}_{i=0}^p$. The corresponding output points are defined as follows:

$$\bar{\mathbf{y}}_j = \begin{cases} \hat{\mathbf{y}}_j, & \text{if } \tau_j \notin \{t_i + \bar{\beta}\}_{i=0}^p, \\ \hat{\mathbf{y}}_j + \boldsymbol{\delta}_j, & \text{if } \tau_j \in \{t_i + \bar{\beta}\}_{i=0}^p, \end{cases}$$

where $\boldsymbol{\delta}_j, j = 1, \dots, m$, are computed using the original experimental data together with an appropriate interpolation technique. Our new objective function is

$$\bar{J}(\alpha, \bar{\beta}) = \sum_{j=1}^m |\mathbf{y}(\bar{\tau}_j | \alpha, \bar{\beta}) - \bar{\mathbf{y}}^j|^2 \approx J(\alpha, \bar{\beta}).$$

The partial derivative of \bar{J} with respect to $\bar{\beta}$ can be determined in a similar manner to the derivation of $\frac{\partial J}{\partial \alpha}$ given above:

$$\begin{aligned} \frac{\partial \bar{J}(\alpha, \bar{\beta})}{\partial \bar{\beta}} &= 2 \sum_{j=1}^m (\mathbf{y}(\bar{\tau}_j | \alpha, \bar{\beta}) - \bar{\mathbf{y}}^j)^\top \frac{\partial \mathbf{g}(\bar{\tau}_j | \alpha, \bar{\beta})}{\partial \mathbf{x}} \boldsymbol{\Gamma}(\bar{\tau}_j | \alpha, \bar{\beta}) \\ &\quad + 2 (\mathbf{y}(\bar{\tau}_j | \alpha, \bar{\beta}) - \bar{\mathbf{y}}^j)^\top \frac{\partial \mathbf{g}(\bar{\tau}_j | \alpha, \bar{\beta})}{\partial \tilde{\mathbf{x}}} \boldsymbol{\Gamma}(\bar{\tau}_j - \alpha | \alpha, \bar{\beta}). \end{aligned} \quad (67)$$

Since it is unlikely that many of the sample times will lie in the set $\{t_i + \bar{\beta}\}_{i=0}^p$, there should be no noticeable difference between minimizing \bar{J} and minimizing J . Indeed, our numerical

results in the next section indicate that this is precisely the case. Thus, during the course of the optimization process, we compute $\frac{\partial J}{\partial \beta}$ using (65) and we approximate $\frac{\partial J}{\partial \beta}$ using (67). This heuristic strategy for determining the cost function and its gradient at a given pair $(\alpha, \bar{\beta})$ is described below.

- Step 1.* Obtain $\mathbf{x}(\cdot|\alpha, \bar{\beta})$, $\mathbf{\Lambda}(\cdot|\alpha, \bar{\beta})$, and $\mathbf{\Gamma}(\cdot|\alpha, \bar{\beta})$ by solving the enlarged time-delay system consisting of the original system (1)-(2) and the auxiliary systems (16)-(17) and (34)-(36).
- Step 2.* Use $\mathbf{x}(\bar{\tau}_j|\alpha, \bar{\beta})$, $j = 1, \dots, m$ to compute $\mathbf{y}(\bar{\tau}_j|\alpha, \bar{\beta})$ through equation (9).
- Step 3.* Use $\mathbf{y}(\bar{\tau}_j|\alpha, \bar{\beta})$, $j = 1, \dots, m$ to compute $J(\alpha, \bar{\beta})$ through equation (4).
- Step 4.* Use $\mathbf{x}(\bar{\tau}_j|\alpha, \bar{\beta})$, $\mathbf{y}(\bar{\tau}_j|\alpha, \bar{\beta})$, $\mathbf{\Lambda}(\bar{\tau}_j|\alpha, \bar{\beta})$, $j = 1, \dots, m$ to compute $\frac{\partial J(\alpha, \bar{\beta})}{\partial \alpha}$ through equation (65).
- Step 5.* Perturb the time points according to (66). Obtain $\mathbf{x}(\bar{\tau}_j|\alpha, \bar{\beta})$, and $\mathbf{\Gamma}(\bar{\tau}_j|\alpha, \bar{\beta})$, $j = 1, \dots, m$ by interpolation, and then calculate $\mathbf{y}(\bar{\tau}_j|\alpha, \bar{\beta})$.
- Step 6.* Use $\mathbf{x}(\bar{\tau}_j|\alpha, \bar{\beta})$, $\mathbf{y}(\bar{\tau}_j|\alpha, \bar{\beta})$, $\mathbf{\Gamma}(\bar{\tau}_j|\alpha, \bar{\beta})$, $j = 1, \dots, m$ to compute $\frac{\partial J(\alpha, \bar{\beta})}{\partial \beta}$ through equation (67).

5 Numerical examples

5.1 Example 1: Zinc Sulphate purification

For our first example, we consider the industrial zinc sulphate purification process described in [25]. In this process, zinc powder is added to a zinc sulphate electrolyte to induce deposition of harmful cobalt and cadmium ions. This is a key step in the production of zinc.

The rates of change of cobalt and cadmium ion concentrations in the electrolyte are described by the following differential equations:

$$V\dot{x}_1(t) = Qx_1^0 - Qx_1(t - \alpha) - c_1u(t - \beta)x_1(t - \alpha) + c_3x_2(t - \alpha), \quad (68)$$

$$V\dot{x}_2(t) = Qx_2^0 - Qx_2(t - \alpha) - c_2v(t)x_2(t - \alpha) + c_4x_1(t - \alpha), \quad (69)$$

and

$$x_1(t) = 3.3 \times 10^{-4}, \quad x_2(t) = 4.0 \times 10^{-3} \quad t \leq 0, \quad (70)$$

where x_1 is the concentration of cobalt ions; x_2 is the concentration of cadmium ions; and u and v are control variables that correspond to the zinc powder reaction surface areas (proportional to the amount of zinc powder added to the reaction tank).

Furthermore, V is the volume of the reaction tank ($V = 400 \text{ m}^3$); Q is the flux of solution ($Q = 200 \text{ m}^3/\text{h}$); c_1 , c_2 , c_3 , c_4 , are model parameters; and x_1^0 and x_2^0 are the concentrations of cobalt and cadmium ions at the inlet of the reaction tank, respectively ($x_1^0 = 6 \times 10^{-4} \text{ g/L}$, $x_2^0 = 9 \times 10^{-3} \text{ g/L}$). Reference [25] considers the parameter identification problem for system

Tab. 1: Control values and switching times for Example 1.

i	1	2	3	4	5	6	7	8
\bar{t}_i	1	2	3	4	5	6	7	8
$\sigma_i \times 10^{-5}$	1.08	1.57	1.24	1.56	1.59	1.43	1.25	1.25
$\bar{\sigma}_i \times 10^{-5}$	5.20	4.70	4.97	4.60	4.53	4.64	4.74	4.62

(68)-(70) with a given state time-delay of $\alpha = 2$. Here, the zinc powder is added with time-delay β . We assume that the system parameters equal to the optimal values reported in [25]:

$$c_1 = 7.828 \times 10^{-4}, \quad c_2 = 2.823 \times 10^{-4}, \quad c_3 = 16.67, \quad c_4 = 7.107 \times 10^2. \quad (71)$$

These values were obtained using data from a real zinc production factory in China. We assume that the terminal time is $T = 8$. We set the input variables u and v as equal to the optimal control functions obtained in [25]:

$$u(t) = \sigma^i, \quad t \in [t_{i-1}, t_i), \quad i = 1, \dots, 8, \quad (72)$$

$$v(t) = \bar{\sigma}^i, \quad t \in [t_{i-1}, t_i), \quad i = 1, \dots, 8, \quad (73)$$

where the values of t_i , σ^i , and $\bar{\sigma}^i$, $k = 1, \dots, 8$ are given constant listed in Table 1. The output of the system is the concentration of cadmium ions:

$$y(t) = x_2(t). \quad (74)$$

Given system (68)-(70) with data (71) and inputs (72)-(73) our goal is to identify the delays α and β . We simulate system (68)-(70) with $[\hat{\alpha}, \hat{\beta}]^\top = [2, 0.25]$ to generate the observed data in Problem (P). The observed data $\hat{y}^j = x_2(\tau_j | \tau, \alpha)$ is sampled at $\tau_j = j/5$, $j = 1, \dots, 40$. Thus, our identification problem is: choose α and β to minimize

$$J(\alpha, \beta) = \sum_{j=1}^{40} |y(\tau_j | \alpha, \beta) - \hat{y}^j|^2 = \sum_{j=1}^{40} |x_2(\tau_j | \alpha, \beta) - x_2(\tau_j | \hat{\alpha}, \hat{\beta})|^2. \quad (75)$$

subject to the dynamic system (68)-(70).

Note that at the optima value, $\tau_j \in I$ happens $q - 1$ times. We solve this problem using a Matlab program that integrate the SQP optimization method with the gradient computation algorithm described in Section 4.

Computational results with different initial guesses are shown in Table 2. The optimization trajectories with initial guess $\alpha = 3$, $\beta = 3$ are displayed in Figure 1. In Table 2 and Figure 1, α^k and β^k are the values of α and β at the k th iteration during the optimization process, while $k = 0$ denotes that of the initial guess. N_1 denotes the times when $t_j + \beta \in \{t_i, i = 1, \dots, 8\}$, $j = 1, \dots, 8$ during the convergence progress. N_2 denotes the times when $\tau_j \in \{t_i + \beta, i = 1, \dots, 8\}$, $j = 1, \dots, 40$ during the convergence progress. We can see from Table 2 and Figure 1 that the optimal trajectory converges to the observed data well with any initial guesses.

Tab. 2: Numerical convergence of the cost values in Example 1.

No.	Initial guess		Cost value at the k th iteration				N_1	N_2
	α^0	β^0	$k = 0$	$k = 5$	$k = 10$	$k = 20$		
1	0.5	0.5	5.865×10^{-5}	9.001×10^{-8}	8.396×10^{-25}	1.151×10^{-27}	0	0
2	1.0	1.0	4.171×10^{-5}	6.265×10^{-8}	2.218×10^{-21}	2.287×10^{-34}	3	0
3	3.0	2.0	9.169×10^{-5}	2.007×10^{-5}	6.624×10^{-7}	8.209×10^{-27}	3	0
4	3.0	3.0	7.828×10^{-5}	2.122×10^{-6}	2.318×10^{-8}	1.141×10^{-26}	6	0

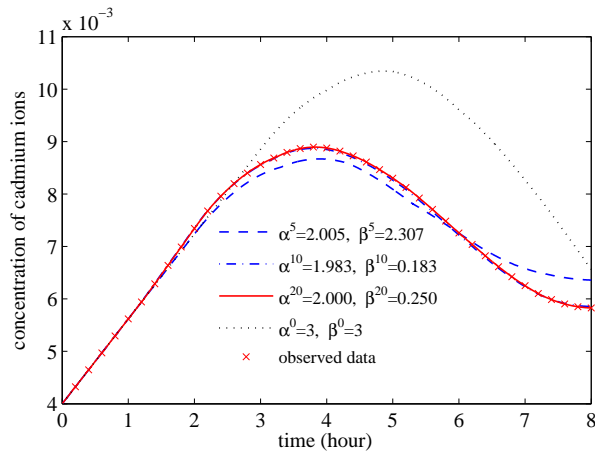


Fig. 1: Numerical convergence of the output trajectory in Example 1 for initial guess No.4.

5.2 Example 2: Sodium aluminate evaporation

We now consider another industrial evaporation process described in [1]. The purpose of this process is to improve the concentration of the mother liquor to reach specific concentration requirement, such that the acid and caustic materials can be re-used. Due to the hysteresis during the solution flowing through the evaporation vessel, the changes of the input solutions can not cause an effect on the changes in the evaporation vessel instantaneously. Thus, there are delays existing in process. For simplicity, we just consider the last two evaporators. The variables that are of interest are the Sodium hydroxide concentration, temperature, and level of the solution in each of the evaporation vessel. Thus, the dynamics in these two evaporation

vessels can be presented by the following differential equations:

$$\frac{dh_1(t)}{dt} = \frac{F_{01}(t - \beta)\rho_0(t - \alpha) + F_2(t - \beta)\rho_2(t - \alpha) - F_1(t)\rho_1(t) + V_0}{A_1\rho_1(t)}, \quad (76)$$

$$\frac{dh_2(t)}{dt} = \frac{F_0(t - \beta)\rho_0(t - \alpha) - F_2(t)\rho_2(t)}{A_1\rho_2(t)}, \quad (77)$$

$$\frac{dC_1(t)}{dt} = \frac{F_{01}(t - \beta)C_0(t - \alpha) + F_2(t - \beta)C_2(t - \alpha) - F_1(t)C_1(t)}{A_1h_1(t)} - \frac{dh_1(t)}{dt} \frac{C_1(t)}{h_1(t)}, \quad (78)$$

$$\frac{dC_2(t)}{dt} = \frac{F_0(t - \beta)C_0(t - \alpha) - F_2(t)C_2(t)}{A_2h_2(t)} - \frac{dh_2(t)}{dt} \frac{C_2(t)}{h_2(t)}. \quad (79)$$

$$\frac{dT_i(t)}{dt} = \frac{\Delta Q_i(t)}{A_i h_i(t) c p_i(t)} - \frac{T_i(t)}{c p_i(t)} \frac{d c p_i}{dt} - \frac{T_i(t) \rho_i(t)}{h_i(t)} \frac{d h_i(t)}{dt}, \quad i = 1, 2, \quad (80)$$

where $i, i = 1, 2$, refer to the i th evaporator; h, T , and C are the state variables representing the level, temperature, and concentration of the solution in the evaporation vessel, respectively; A is the cross-sectional area of the evaporation vessel; F is the flow rate of the product solution; V_0 is the amount of vapor from other heat sources mixed with the solution; F_0 is the flow rate of the feed; C_0 is the concentration of the feed; $\Delta Q_i, i = 1, 2$, are the heat changes in the evaporation vessel depends on the live steam flow rate; cp and ρ are the specific heat capacity and the density of the solution, respectively, which depend on the concentration and temperature. cp, ρ , and ΔQ are calculated by using the formulas given in [1].

The initial conditions for system (76)-(80) are

$$\begin{aligned} \mathbf{x}(t) &= [h_1(t), h_2(t), C_1(t), C_2(t), T_1(t), T_2(t)]^\top \\ &= [1.91, 110.6, 73.0, 1.91, 97.2, 54.5]^\top, \quad t \leq 0. \end{aligned} \quad (81)$$

Here, the inputs are $\mathbf{u} = [u_1, \dots, u_5]^\top = [F_1, F_2, F_0, F_{01}, V]^\top$, where V denotes the live steam flow rate. α is an unknown state-delays, β is an unknown input-delay. Assume that the terminal time of this system is $T = 240$ minutes. The input function is

$$u_k(t) = \sigma_k^i, \quad t \in [t_{i-1}, t_i], \quad k = 1, 2, 3, 5, \quad i = 1, \dots, 24, \quad (82)$$

and

$$u_4(t) = 0.165, \quad t \in [0, 240], \quad (83)$$

where $\sigma_k^i, i = 1, \dots, 24$, are given constant whose heights are shown in the Figure 2. The output is

$$\mathbf{y}(t) = [C_1(t), C_2(t)]^\top, \quad t \leq 240.$$

We use the output trajectory of (76)-(81) with $[\hat{\alpha}, \hat{\beta}] = [15, 6]^\top$ to generate the observed data for Problem (P). We set

$$\hat{\mathbf{y}}^j = [x_2(\tau_j | \hat{\alpha}, \hat{\beta}), x_5(\tau_j | \hat{\alpha}, \hat{\beta})]^\top, \quad j = 1, \dots, 24.$$

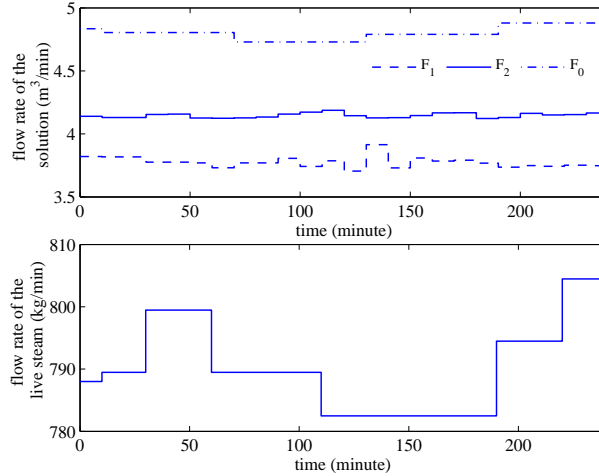


Fig. 2: Inputs of Example 2

where, $\tau_{j+1} - \tau_j = 10$, Thus, our identification problem is: choose α and β to minimize

$$\begin{aligned}
 J(\alpha, \beta) &= \sum_{j=1}^{24} |\mathbf{y}(\tau_j | \alpha, \beta) - \hat{\mathbf{y}}^j|^2 \\
 &= \sum_{j=1}^{24} |x_2(\tau_j | \alpha, \beta) - x_2(\tau_j | \hat{\alpha}, \hat{\beta})|^2 + |x_5(\tau_j | \alpha, \beta) - x_5(\tau_j | \hat{\alpha}, \hat{\beta})|^2
 \end{aligned} \quad (84)$$

subject to the dynamic (76)-(81).

We solved this problem using a Matlab program that integrates the SQP optimization method with the gradient computation algorithm described in Section 4. The convergence progress of the program is shown in Table 3 for four sets of initial guesses. The convergence progress corresponding to the initial guess $\alpha = 36$, $\beta = 36$, is shown in Figure 3. In Table 3 and Figure 3, α^l and β^l are the values of α and β at the l th iteration, while $l = 0$ denotes that of the initial guess. N_1 denotes the times when $t_j + \beta \in \{t_i, i = 1, \dots, 24\}$, $j = 1, \dots, 24$, during the convergence progress. N_2 denotes the times when $\tau_j \in \{t_i + \beta, i = 1, \dots, 24\}$, $j = 1, \dots, 24$ during the convergence progress. We can see that the optimization results converge from the initial guesses to the optimal solution.

6 Conclusion

In this paper, we develop a gradient-based computational method for solving a time delay identification problem, while the input function of the nonlinear time-delay system is piecewise-constant. We assume that the time-delays—one involving the state variables and the other involving the input variables—are unknown and need to be estimated using experimental data. This method is unified in the sense that the gradient of the cost function with respect to the input delay is obtained by a auxiliary delay-differential system from $t = 0$ to $t = T$ with jump conditions at the delayed control switching time points. The industrial examples demonstrate

Tab. 3: Numerical convergence of the cost values in Example 2.

No.	Initial guess		Cost values at the i th iteration				N_1	N_2
	α^0	β^0	$l = 0$	$l = 5$	$l = 10$	$l = 20$		
1	12	12	0.077	4.462×10^{-3}	2.782×10^{-8}	2.446×10^{-8}	1	0
2	24	24	0.559	4.496×10^{-4}	3.259×10^{-7}	4.684×10^{-8}	2	0
3	30	30	0.939	5.300×10^{-3}	3.259×10^{-7}	4.524×10^{-8}	1	0
4	36	36	1.170	8.751×10^{-1}	1.728×10^{-4}	8.455×10^{-8}	0	0
5	48	48	1.628	8.862×10^{-4}	6.432×10^{-7}	4.802×10^{-8}	0	0

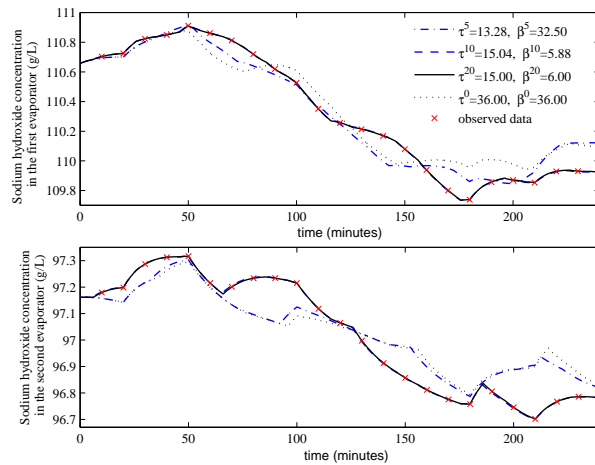


Fig. 3: Numerical convergence of the output trajectory in Example 2 for initial guess No.4.

that this approach is highly effective. In particular, it converges quickly even when the initial estimates for the delays and parameters are far away from the optimal values.

References

- [1] Q. Q. Chai, C. H. Yang, K. L. Teo, W. H. Gui, Time-delayed optimal control of an industrial-scale evaporation process sodium aluminate solution. *Control Engineering Practice*, 2012,20 (6): 618-628.
- [2] S. Dimartino, C. Boi, G. C. Sarti, A validated model for the simulation of protein purification through affinity membrane chromatography, *Journal of Chromatography A*, 2011, 1218(13): 1677-1690.
- [3] G. Fernholz, S. Engell, L. U. Kreul, A. Gorak, Optimal operation of a semi-batch reactive distillation column. *Computers and Chemical Engineering*, 2000, 24(2): 1569-1575.
- [4] L. Wang, W. Gui, K. L. Teo, R. Loxton, C. Yang, Time delayed optimal control problems with multiple characteristic time points: Computation and industrial applications. *Journal of Industrial and Management Optimization*, 2009, 5: 705-718.
- [5] K. L. Teo, C. J. Goh, K. H. Wong, *A Unified Computational Approach to Optimal Control Problems*, Longman Scientific and Technical: Essex, U. K., 1991.
- [6] R. C. Loxton, K. L. Teo, V. Rehbock, K. F. C. Yiu, Optimal control problems with a continuous inequality constraint on the state and the control, *Automatica*, 2009, 45(10): 2250-2257.
- [7] H. R. Marzban, S. M. Hoseini, Solution of linear optimal control problems with time delay using a composite Chebyshev finite difference method, *Optimal Control Applications and Methods*, (2012): DOI: 10.1002/oca.2019.
- [8] Richard2003 J. P. Richard2003, Time-delay systems: An overview of some recent advances and open problems, *Automatica*, 2003, 39(10): 1667-1694.
- [9] M. Anguelova, B. Wennberg, State elimination and identifiability of the delay parameter for nonlinear time-delay systems, *Automatica* 44 (2008) 1373-1378.
- [10] L. Belkoura, J. P. Richard2003, M. Fliess, Parameters estimation of systems with delayed and structured entries, *Automatica*, 2009, 45(4): 1117-1125.
- [11] B. Ni, D. Xiao, S. L. Shah, Time delay estimation for MIMO dynamical systems—With time-frequency domain analysis, *Journal of Process Control* 20 (2010) 83-94.
- [12] M. Liu, Q. G. Wang, B. Huang, C. C. Hang, Improved identification of continuous-time delay processes from piecewise step tests. *Journal of Process Control*, 2007, 17(1): 51-57.

-
- [13] R. Loxton, K. L. Teo, V. Rehbock, An optimization approach to state-delay identification, *IEEE Transactions on Automatic Control*, 2010, 55(9): 2113-2119.
- [14] M. S. Charalambides, N. Shah, C. C. Pantelides, Synthesis of batch reaction/distillation processes using detailed dynamic models, *Computers and Chemical Engineering*, 1995, 19(11): 167-174.
- [15] Q. Lin, R. Loxton, K. L. Teo, Y. H. Wu, A New Computational Method for Optimizing Nonlinear Impulsive Systems. *Dynamics of Continuous, Discrete and Impulsive Systems B*. 2011, 18(1): 59-76.
- [16] R. Loxton, K. L. Teo, V. Rehbock, Computational Method for a Class of Switched System Optimal Control Problems. *IEEE Transactions on Automatic Control*, 2009, 54(10): 2455-2460.
- [17] U. N. Ahmed, *Dynamic Systems and Control with Applications*, Singapore: World Scientific, 2006.
- [18] C. Seatzu, D. Corona, A. Giua, A. Bemporad, Optimal control of continuous-time switched affine systems. *IEEE Transactions on Automatic Control*, 2006, 51(5): 726-741.
- [19] M. Kamgarpour, C. Tomlin, On optimal control of non-autonomous switched systems with a fixed mode sequence, *Automatica*, 2012, 48(6), 1177-1181.
- [20] C. Liu, Z. Gong, E. Feng, H. Yin, Modelling and optimal control for nonlinear multistage dynamical system of microbial fed-batch culture, *Journal of Industrial and Management Optimization*, 2009, 5(4): 835-850.
- [21] I. Mallocci, J. Daafouz, C. Iung, Stability and stabilization of two time scale switched systems in discrete time, *IEEE Transactions on Automatic Control*, 2010, 55(6):1434-1438.
- [22] C. Jiang, K. L. Teo, R. Loxton, G. R. Duan, Optimal switched impulsive control problems with large perturbations, *International Journal of Innovative Computing, Information and Control*, 2012, 8(9): 6235-6257.
- [23] T. Ruby, V. Rehbock, Numerical solutions of optimal switching control problems. *Optimization and Control with Applications*, 2005, 96:447-459.
- [24] X. Xu, P. J. Antsaklis, Optimal control of switched systems based on parameterization of the switching instants, *IEEE Transactions on Automatic Control*, 2004, 49(1): 2-16.
- [25] L. Y. Wang, W. H. Gui, K. L. Teo, R. Loxton, C. H. Yang, Optimal control problems arising in the zinc sulphate electrolyte purification process, *Journal of Global Optimization*, 2012, 54(2): 307-323.