# Time Delayed Optimal Control Problems with Multiple Characteristic Time Points: Computation and Industrial Applications

Ling Yun Wang[1,2], Wei Hua Gui[1], Kok Lay Teo[2]*, Ryan C. Loxton[2], Chun Hua Yang[1]

1. School of Information Science and Engineering,

Central South University, Changsha, China, 410083

2. Department of Mathematics and Statistics,

Curtin University of Technology, Perth, Australia, 6105

**Abstract:** In this paper, we consider a class of optimal control problems involving time delayed dynamical systems and subject to continuous state inequality constraints. We show that this type of problem can be approximated by a sequence of time delayed optimal control problems subject to inequality constraints in canonical form and with multiple characteristic time points appearing in the cost and constraint functions. We derive formulae for the gradient of the cost and constraint functions of the approximate problems. On this basis, each approximate problem can be solved using a gradient-based optimization technique. The computational method obtained is then applied to an industrial problem arising in the study of purification process of zinc sulphate electrolyte. The results are highly satisfactory.

**Keywords:** Time delayed system; Multiple characteristic time points; Continuous state inequality constraints; Optimal control; Zinc sulphate electrolyte; Purification process

## 1 Introduction

Time delay exists inevitably in many practical control systems, and hence there have been intensive investigation amongst researchers working on control systems with time delay in numerous research fields such as aeronautical, astronautical, mechanical, chemical and electrical engineering. Many methodologies have been proposed to deal with

*Corresponding author: k.l.teo@curtin.edu.au

the control and optimal control problems of systems with time delay, see, for example, in [1–9], respectively. The methods proposed in [2, 3] are for linear dynamical systems and without constraints on the state and control variables. A novel optimal control method for linear systems with time delay through a particular transformation proposed in [5] in vibration control is developed on the basis of [4]. For the methods proposed in [6, 7], there is also no continuous constraint on the state and control variables. The first attempt to use iterative dynamic programming with time-delay systems was made by Dadebo and Luus (see, [8]), who used piecewise constant control policy. The Taylor series expansion for the delay terms was used in [9] to convert the given system into a non-delay system for which the optimal control policy can be readily established. The class of optimal control problems considered in [10] is more general. It involves both equality and inequality constraints on the state and control variables in canonical form. However, only a single characteristic time point is allowed in the cost as well as the constraint functions. Thus, it is not applicable to optimal control problems involving system parameters and/or control in the cost and constraint functions, where these cost and constraint functions also contain multiple characteristic times. Such problems arise naturally in the study of control of chemical reaction systems involving the purification process of zinc sulphate electrolyte. There are two parts associated with such an application. A mathematical model is first constructed. This model is described by a system of time delayed ordinary differential equations with some system parameters are yet to be specified. To identify the values of these parameters, measurements are carried out at a set of discrete time points. Then, the problem of identifying these parameter values is equivalent to that of finding the parameter values such that the solution of the time delayed differential system will "best" fit the observed data at the observation time points. This problem can be formulated as an optimal parameter selection problem with multiple characteristic time points. Loxton, Teo and Rehbock (see, [13]) also solve the optimal control problem with multiple characteristic time points, but it does not solve the system with time delay. For relevant references for optimal parameter selection problems, see, for example, [11,12]. Now, with the parameters being best identified, the problem of minimizing the use of zinc powder, which is used to remove the impurities metallic ions in the zinc sulphate electrolyte, can be formulated as an optimal control problem involving the same system of time delayed ordinary differential equations subject to continuous state constraints.

In response to the motivation given above, we consider a class of time delayed optimal control problems involving system parameters and control in the cost and constraint functions and subject to continuous state inequality constraints, where these cost and constraint functions also contain multiple characteristic times. We shall develop a computational method to solve this class of time delayed optimal control problems. The method obtained will then be applied to the study of the chemical reaction system involving purification process of zinc sulphate electrolyte mentioned above.

## 2 Problem statement

Consider a process described by the following system of time delayed differential equations defined on the fixed time interval $(0, T]$:

$$\frac{\mathrm{d}\boldsymbol{x}(t)}{\mathrm{d}t} = \boldsymbol{f}(t, \boldsymbol{x}(t), \boldsymbol{x}(t-h), \boldsymbol{u}(t), \boldsymbol{\zeta}), \tag{2.1}$$

where $\boldsymbol{x} = [x_1, \ldots, x_n]^\top \in \mathbb{R}^n$, $\boldsymbol{u} = [u_1, \ldots, u_r]^\top \in \mathbb{R}^r$, $\boldsymbol{\zeta} = [\zeta_1, \ldots, \zeta_s]^\top \in \mathbb{R}^s$ are, respectively, the state, control and system parameter vectors; $\boldsymbol{f} = [f_1, \ldots, f_n]^\top \in \mathbb{R}^n$; $h$ is the time delay satisfying $0 < h < T$; and the superscript $\top$ denotes transpose. Here, only the case of a single time delay is considered. However, all the results obtained can be easily extended to the case of multiple time delays. The initial function for the differential equation (2.1) is

$$\boldsymbol{x}(t) = \boldsymbol{\phi}(t), \quad t \in [-h, 0), \tag{2.2a}$$

$$\boldsymbol{x}(0) = \boldsymbol{x}^0, \tag{2.2b}$$

where $\boldsymbol{\phi}(t) = [\phi_1(t), \ldots, \phi_n(t)]^\top$ is a given piecewise continuous function from $[-h, 0)$ into $\mathbb{R}^n$, and $\boldsymbol{x}^0$ is a given vector in $\mathbb{R}^n$.

Let $t_k$, $k = 1, \ldots, N$, be given time points in $[0, T]$. We assume that the control takes the form given below:

$$\boldsymbol{u}(t) = \boldsymbol{\gamma}(t), \quad t \in [-h, 0), \tag{2.2c}$$

$$\boldsymbol{u}(t) = \sum_{k=1}^{N} \boldsymbol{\sigma}^k \chi_{[t_{k-1}, t_k)}(t), \quad t \in [0, T], \tag{2.3}$$

where $\boldsymbol{\gamma}(t) = [\gamma_1(t), \ldots, \gamma_r(t)]^\top$ is a given piecewise continuous function from $[-h, 0)$ into $\mathbb{R}^r$, while $\chi_{[t_{k-1}, t_k)}$ denotes the indicator function of the interval $[t_{k-1}, t_k)$ defined by

$$\chi_I(t) = \begin{cases} 1, & \text{if} \quad t \in I, \\ 0, & \text{otherwise.} \end{cases} \tag{2.4}$$

Define $\boldsymbol{\sigma}^k = [\sigma_1^k, \ldots, \sigma_r^k]^\top$ and $\boldsymbol{\sigma} = [(\boldsymbol{\sigma}^1)^\top, \ldots, (\boldsymbol{\sigma}^N)^\top]^\top$. Let $U \subset \mathbb{R}^r$ be both compact and convex. A function $\boldsymbol{u}$ given by (2.3) with $\boldsymbol{\sigma}^k \in U$, $k = 1, \ldots, N$, is called an admissible control. Let $\mathcal{U}$ be the set of all such admissible controls, and let $\Theta$ be the set containing all $\boldsymbol{\sigma} = [(\boldsymbol{\sigma}^1)^\top, \ldots, (\boldsymbol{\sigma}^N)^\top]^\top$ with $\boldsymbol{\sigma}^k \in U$, $k = 1, \ldots, N$. Clearly, each $\boldsymbol{u} \in \mathcal{U}$ corresponds uniquely to a $\boldsymbol{\sigma} \in \Theta$ and vice versa. Let $Z \subset \mathbb{R}^s$ be both compact and convex, and we assume that $\boldsymbol{\zeta} = [\zeta_1, \ldots, \zeta_s]^\top \in Z$. With $\boldsymbol{u} \in \mathcal{U}$, the system of time delayed differential equations (2.1) can be written as:

$$\frac{\mathrm{d}\boldsymbol{x}(t)}{\mathrm{d}t} = \tilde{\boldsymbol{f}}(t, \boldsymbol{x}(t), \boldsymbol{x}(t-h), \boldsymbol{\sigma}, \boldsymbol{\zeta}), \tag{2.5}$$

where $\tilde{\boldsymbol{f}}(t, \boldsymbol{x}(t), \boldsymbol{x}(t-h), \boldsymbol{\sigma}, \boldsymbol{\zeta}) = \boldsymbol{f}(t, \boldsymbol{x}(t), \boldsymbol{x}(t-h), \boldsymbol{u}(t), \boldsymbol{\zeta})$ with $\boldsymbol{u}$ taking the form of (2.3).

For each $(\boldsymbol{\sigma}, \boldsymbol{\zeta}) \in \Theta \times Z$, let $\boldsymbol{x}(\cdot \mid \boldsymbol{\sigma}, \boldsymbol{\zeta})$ be the corresponding vector-valued function which is absolutely continuous on $(0, T]$ and satisfies the differential equation (2.5) almost everywhere on $(0, T]$ and the initial conditions (2.2). This function is called the solution of system (2.5) with initial conditions (2.2) corresponding to the parameter vector $(\boldsymbol{\sigma}, \boldsymbol{\zeta}) \in \Theta \times Z$.

We may now state a time delayed optimal control problem with multiple characteristic time points appearing in the cost and constraint functions as follows:

**Problem (P1)**. Given system (2.5) with initial conditions (2.2), find a parameter vector $(\boldsymbol{\sigma}, \boldsymbol{\zeta}) \in \Theta \times Z$ such that the cost function

$$\begin{aligned} g_0(\boldsymbol{\sigma}, \boldsymbol{\zeta}) =& \Phi_0(\boldsymbol{x}(\tau_1 \mid \boldsymbol{\sigma}, \boldsymbol{\zeta}), \ldots, \boldsymbol{x}(\tau_{M+1} \mid \boldsymbol{\sigma}, \boldsymbol{\zeta})) \\ &+ \int_0^T \tilde{\mathcal{L}}_0(t, \boldsymbol{x}(t \mid \boldsymbol{\sigma}, \boldsymbol{\zeta}), \boldsymbol{x}(t-h \mid \boldsymbol{\sigma}, \boldsymbol{\zeta}), \boldsymbol{\sigma}, \boldsymbol{\zeta}) \mathrm{d}t \end{aligned} \tag{2.6}$$

is minimized subject to the following canonical inequality constraints:

$$\begin{aligned} g_m(\boldsymbol{\sigma}, \boldsymbol{\zeta}) =& \Phi_m(\boldsymbol{x}(\tau_1 \mid \boldsymbol{\sigma}, \boldsymbol{\zeta}), \ldots, \boldsymbol{x}(\tau_{M+1} \mid \boldsymbol{\sigma}, \boldsymbol{\zeta})) \\ &+ \int_0^T \tilde{\mathcal{L}}_m(t, \boldsymbol{x}(t \mid \boldsymbol{\sigma}, \boldsymbol{\zeta}), \boldsymbol{x}(t-h \mid \boldsymbol{\sigma}, \boldsymbol{\zeta}), \boldsymbol{\sigma}, \boldsymbol{\zeta}) \mathrm{d}t \geq 0, \quad m = 1, \ldots, N, \end{aligned} \tag{2.7}$$

where the time points $\tau_i$, $0 < \tau_i < T$, $i = 1, \ldots, M$, are referred to as the characteristic time points. We use the convention that $\tau_0 = 0$ and $\tau_{M+1} = T$. Here, with $\boldsymbol{u} \in \mathcal{U}$,

$$\begin{aligned} \tilde{\mathcal{L}}_m(t, \boldsymbol{x}(t \mid \boldsymbol{\sigma}, \boldsymbol{\zeta}), \boldsymbol{x}(t-h \mid \boldsymbol{\sigma}, \boldsymbol{\zeta}), \boldsymbol{\sigma}, \boldsymbol{\zeta}) = \mathcal{L}_m(t, \boldsymbol{x}(t \mid \boldsymbol{\sigma}, \boldsymbol{\zeta}), \boldsymbol{x}(t-h \mid \boldsymbol{\sigma}, \boldsymbol{\zeta}), \boldsymbol{u}(t), \boldsymbol{\zeta}), \\ m = 0, 1, \ldots, N. \end{aligned} \tag{2.8}$$

In a standard optimal control problem, each canonical constraint (including the cost function which corresponds to $m = 0$) depends only on one characteristic time point. However, for Problem (P1), there are multiple characteristic time points appearing in the cost and constraint functions.

We assume that the following conditions are satisfied.

(A1) $\boldsymbol{f} : [0, T] \times \mathbb{R}^{2n+r+s} \to \mathbb{R}^n$; $\boldsymbol{\phi} : [-h, 0) \to \mathbb{R}^n$;
$\Phi_m : \mathbb{R}^{n+\cdots+n} \to \mathbb{R}$, $m = 0, 1, \ldots, N$;
$\mathcal{L}_m : [0, T] \times \mathbb{R}^{2n+r+s} \to \mathbb{R}$, $m = 0, 1, \ldots, N$.

(A2) For each compact subset $V \times W \subset \mathbb{R}^{r+s}$, there exists a positive constant $K$ such that, for all $(t, \boldsymbol{x}, \boldsymbol{y}, \boldsymbol{u}, \boldsymbol{\zeta}) \in [0, T] \times \mathbb{R}^{2n} \times V \times W$, $|\boldsymbol{f}(t, \boldsymbol{x}, \boldsymbol{y}, \boldsymbol{u}, \boldsymbol{\zeta})| \leq K(1 + |\boldsymbol{x}| + |\boldsymbol{y}|)$, $|\mathcal{L}_m(t, \boldsymbol{x}, \boldsymbol{y}, \boldsymbol{u}, \boldsymbol{\zeta})| \leq K(1 + |\boldsymbol{x}| + |\boldsymbol{y}|)$, $m = 0, 1, \ldots, N$.

(A3) $\boldsymbol{f}(t, \boldsymbol{x}, \boldsymbol{y}, \boldsymbol{u}, \boldsymbol{\zeta})$ and $\mathcal{L}_m(t, \boldsymbol{x}, \boldsymbol{y}, \boldsymbol{u}, \boldsymbol{\zeta})$, $m = 0, 1, \ldots, N$, are piecewise continuous on $(0, T]$ for each $(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{u}, \boldsymbol{\zeta}) \in \mathbb{R}^{2n+r+s}$, and continuously differentiable with respect to each of the components of $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{u}$ and $\boldsymbol{\zeta}$ for each $t \in [0, T]$.

(A4) $\Phi_m$, $m = 0, 1, \ldots, N$, are continuously differentiable on $\mathbb{R}^{n+\cdots+n}$, $\boldsymbol{\phi}$ and $\boldsymbol{\gamma}$ are piecewise continuous on $[-h, 0)$.

**Remark 2.1.** For each $(\boldsymbol{\sigma}, \boldsymbol{\zeta}) \in \mathbb{R}^r \times \mathbb{R}^s$ (and hence $(\boldsymbol{\sigma}, \boldsymbol{\zeta}) \in \Theta \times Z$), there exists a unique absolutely continuous vector-valued function $\boldsymbol{x}(\cdot \mid \boldsymbol{\sigma}, \boldsymbol{\zeta})$ which satisfies the system (2.5) almost everywhere on $(0, T]$ and the initial conditions (2.2).

# 3 Continuous state inequality constraints

In practice, physical constraints are often not expressed by canonical constraints defined by (2.7). Instead, they are described by continuous state inequality constraints given by

$$h_i(t, \boldsymbol{x}(t \mid \boldsymbol{\sigma}, \boldsymbol{\zeta}), \boldsymbol{\zeta}) \geq 0, \quad t \in [0, T], \quad i = 1, \ldots, N, \tag{3.1}$$

where $h_i$, $i = 1, \ldots, N$, are real valued functions defined on $[0, T] \times \mathbb{R}^{n+s}$.

It is assumed that the following conditions are satisfied.

(A5) For each $i = 1, \ldots, N$, $h_i : [0, T] \times \mathbb{R}^{n+s} \to \mathbb{R}$ is continuously differentiable.

(A6) For any $(\boldsymbol{\sigma}, \boldsymbol{\zeta}) \in \Theta \times Z$, there exists a $(\bar{\boldsymbol{\sigma}}, \bar{\boldsymbol{\zeta}}) \in (\Theta \times Z)^\circ$, where $(\Theta \times Z)^\circ$ denotes the interior of $(\Theta \times Z)$, such that $\eta(\bar{\boldsymbol{\sigma}}, \bar{\boldsymbol{\zeta}}) + (1 - \eta)(\boldsymbol{\sigma}, \boldsymbol{\zeta}) \in (\Theta \times Z)^\circ$, $\forall\, \eta \in (0, 1]$.

We may now describe the corresponding optimal control problem as follows.

**Problem (P2)**. Subject to the dynamical system (2.5) with initial conditions (2.2), find a $(\boldsymbol{\sigma}, \boldsymbol{\zeta}) \in \Theta \times Z$ such that the cost function

$$g_0(\boldsymbol{\sigma}, \boldsymbol{\zeta}) = \Phi_0(\boldsymbol{x}(T \mid \boldsymbol{\sigma}, \boldsymbol{\zeta})) + \int_0^T \tilde{\mathcal{L}}_0(t, \boldsymbol{x}(t \mid \boldsymbol{\sigma}, \boldsymbol{\zeta}), \boldsymbol{x}(t-h \mid \boldsymbol{\sigma}, \boldsymbol{\zeta}), \boldsymbol{\sigma}, \boldsymbol{\zeta}) \mathrm{d}t \tag{3.2}$$

is minimized subject to the continuous state inequality constraints (3.1), where with $\boldsymbol{u} \in \mathcal{U}$,

$$\tilde{\mathcal{L}}_0(t, \boldsymbol{x}(t \mid \boldsymbol{\sigma}, \boldsymbol{\zeta}), \boldsymbol{x}(t-h \mid \boldsymbol{\sigma}, \boldsymbol{\zeta}), \boldsymbol{\sigma}, \boldsymbol{\zeta}) = \mathcal{L}_0(t, \boldsymbol{x}(t \mid \boldsymbol{\sigma}, \boldsymbol{\zeta}), \boldsymbol{x}(t-h \mid \boldsymbol{\sigma}, \boldsymbol{\zeta}), \boldsymbol{u}(t), \boldsymbol{\zeta}).$$

Problem (P2) cannot be solved as such due to the presence of the continuous state inequality constraints. However, these continuous inequality constraints can be approximated by a sequence of inequality constraints in canonical form using the constraint transcription method introduced in [16]. Details are given below.

For each $i = 1, \ldots, N$, the continuous state inequality constraint (3.1) is equivalent to

$$g_i(\boldsymbol{\sigma}, \boldsymbol{\zeta}) = \int_0^T \min\{h_i(t, \boldsymbol{x}(t \mid \boldsymbol{\sigma}, \boldsymbol{\zeta}), \boldsymbol{\zeta}), 0\} \mathrm{d}t = 0. \tag{3.3}$$

However, the equality constraint (3.3) is non-differentiable at the points when $h_i = 0$. We replace $\min\{h_i(t, \boldsymbol{x}(t \mid \boldsymbol{\sigma}, \boldsymbol{\zeta}), \boldsymbol{\zeta}), 0\}$ by $\mathcal{L}_{i,\varepsilon}(t, \boldsymbol{x}(t \mid \boldsymbol{\sigma}, \boldsymbol{\zeta}), \boldsymbol{\zeta})$, where

$$\mathcal{L}_{i,\varepsilon}(t, \boldsymbol{x}(t \mid \boldsymbol{\sigma}, \boldsymbol{\zeta}), \boldsymbol{\zeta}) = \begin{cases} h_i(t, \boldsymbol{x}(t \mid \boldsymbol{\sigma}, \boldsymbol{\zeta}), \boldsymbol{\zeta}), & \text{if } \quad h_i(t, \boldsymbol{x}(t \mid \boldsymbol{\sigma}, \boldsymbol{\zeta}), \boldsymbol{\zeta}) < -\varepsilon \\ -(h_i(t, \boldsymbol{x}(t \mid \boldsymbol{\sigma}, \boldsymbol{\zeta}), \boldsymbol{\zeta}) - \varepsilon)^2/4\varepsilon, & \text{if} -\varepsilon \le h_i(t, \boldsymbol{x}(t \mid \boldsymbol{\sigma}, \boldsymbol{\zeta}), \boldsymbol{\zeta}) \le \varepsilon \\ 0, & \text{if } \quad h_i(t, \boldsymbol{x}(t \mid \boldsymbol{\sigma}, \boldsymbol{\zeta}), \boldsymbol{\zeta}) > \varepsilon \end{cases} \tag{3.4}$$

while $\varepsilon > 0$ is an adjustable parameter controlling the accuracy of the approximation. Now, for each $i = 1, \dots, N$, the continuous state inequality constraint (3.1) is approximated by

$$g_{i,\varepsilon}(\boldsymbol{\sigma}, \boldsymbol{\zeta}) = \gamma + \int_0^T \mathcal{L}_{i,\varepsilon}(t, \boldsymbol{x}(t \mid \boldsymbol{\sigma}, \boldsymbol{\zeta}), \boldsymbol{\zeta}) \mathrm{d}t \ge 0, \tag{3.5}$$

where $\gamma > 0$ is an adjustable parameter controlling the feasibility of the constraint (3.1). Thus, Problem (P2) is approximated by a sequence of optimal control problems.
**Problem (P2($\varepsilon, \gamma$)).** Given the time delayed system (2.5) with initial conditions (2.2), find a $(\boldsymbol{\sigma}, \boldsymbol{\zeta}) \in \Theta \times Z$ such that the cost function (3.2) is minimized over $\Theta \times Z$ subject to the canonical inequality constraint (3.5).

Clearly, for each $\varepsilon$ and $\gamma$, Problem (P2($\varepsilon, \gamma$)) is in the form of Problem (P1).

The following theorem shows that for any $\varepsilon > 0$, if $\gamma$ is chosen sufficiently small, the solution of the corresponding optimal Problem (P2($\varepsilon, \gamma$)) will satisfy the continuous state inequality constraint (3.1).
**Theorem 3.1.** For each $\varepsilon > 0$, there exists a $\gamma(\varepsilon) > 0$ such that if (3.5) with $\gamma < \gamma(\varepsilon)$ are satisfied for some $(\boldsymbol{\sigma}, \boldsymbol{\zeta}) \in \Theta \times Z$, then the original constraints (3.1) are also satisfied at $(\boldsymbol{\sigma}, \boldsymbol{\zeta}) \in \Theta \times Z$.

The proof of Theorem 3.1 can be found in Chapter 8 of [16]. On the basis of this theorem, Problem (P2) can be solved through solving a sequence of Problems (P2($\varepsilon, \gamma$)), each of which is a special case of Problem (P1). Thus, it suffices to develop a computational method for Problem (P1).

# 4 A Computational Procedure

Problem (P1) can be viewed as a mathematical programming problem. The cost function (2.6) is to be minimized subject to the bound constraints on the parameter vector and the canonical constraints (2.7). To calculate the values of the cost and constraint functions, as well as their gradients with respect to each parameter vector, the dynamical system (2.5) is used implicitly in the calculation of the values of the cost and constraint functions for each parameter vector. That is, for each parameter vector, dynamical system (2.5) is solved, giving rise to the corresponding solution. This solution is then used together with the given parameter vector for the calculation of the values of the cost and constraint functions. Thus, for each $\varepsilon$ and $\gamma$, Problem (P2($\varepsilon, \gamma$)) can be solved by any gradient-based optimization method, such as the sequential quadratic programming approximation scheme with active set strategy (see, for example, [14]) provided the gradients of the cost and constraint functions with respect to each given parameter vector can be calculated. This task will be much more involved. They require not only the solution of the dynamical system, but also the solution of respective auxiliary dynamical system known as the co-state system for each of these cost and constraint functions. The derivations of these gradients are given below.

For each $m = 0, 1, \ldots, N$, consider the following system, which is known as the corresponding co-state system:

$$\frac{\mathrm{d}(\boldsymbol{\lambda}^m(t))^\top}{\mathrm{d}t} = -\frac{\partial H_m(t, \boldsymbol{x}(t), \boldsymbol{y}(t), \boldsymbol{\sigma}, \boldsymbol{\zeta}, \boldsymbol{\lambda}^m(t))}{\partial \boldsymbol{x}} - \frac{\partial \hat{H}_m(t, \boldsymbol{z}(t), \boldsymbol{x}(t), \boldsymbol{\sigma}, \boldsymbol{\zeta}, \hat{\boldsymbol{\lambda}}^m(t))}{\partial \boldsymbol{x}} \tag{4.1a}$$

where $t \in (\tau_{k-1}, \tau_k)$, $k = 1, \ldots, M+1$, with the jump conditions:

$$\boldsymbol{\lambda}^m(\tau_k^+)^\top - \boldsymbol{\lambda}^m(\tau_k^-)^\top = -\frac{\partial \Phi_m(\boldsymbol{x}(\tau_1), \ldots, \boldsymbol{x}(\tau_{M+1}))}{\partial \boldsymbol{x}(\tau_k)}, \quad \text{for} \quad k = 1, \ldots, M, \tag{4.1b}$$

and the terminal condition

$$(\boldsymbol{\lambda}^m(T))^\top = \frac{\partial \Phi_m(\boldsymbol{x}(\tau_1), \ldots, \boldsymbol{x}(\tau_M), \boldsymbol{x}(T))}{\partial \boldsymbol{x}(T)}, \tag{4.1c}$$

$$\boldsymbol{\lambda}^m(t) = 0, \quad t > T, \tag{4.1d}$$

where

$$\boldsymbol{y}(t) = \boldsymbol{x}(t - h \mid \boldsymbol{\sigma}, \boldsymbol{\zeta}), \tag{4.1e}$$

$$\boldsymbol{z}(t) = \boldsymbol{x}(t + h \mid \boldsymbol{\sigma}, \boldsymbol{\zeta}), \tag{4.1f}$$

$$\hat{\boldsymbol{\lambda}}^m(t) = \boldsymbol{\lambda}^m(t + h), \tag{4.1g}$$

$$H_m(t, \boldsymbol{x}, \boldsymbol{y}, \boldsymbol{\sigma}, \boldsymbol{\zeta}, \boldsymbol{\lambda}^m) = \tilde{\mathcal{L}}_m(t, \boldsymbol{x}, \boldsymbol{y}, \boldsymbol{\sigma}, \boldsymbol{\zeta}) + (\boldsymbol{\lambda}^{\boldsymbol{m}})^\top \tilde{\boldsymbol{f}}(t, \boldsymbol{x}, \boldsymbol{y}, \boldsymbol{\sigma}, \boldsymbol{\zeta}), \tag{4.2a}$$

$$\begin{aligned} \hat{H}_m(t, \boldsymbol{z}, \boldsymbol{x}, \boldsymbol{\sigma}, \boldsymbol{\zeta}, \hat{\boldsymbol{\lambda}}^m) =& \tilde{\mathcal{L}}_m(t + h, \boldsymbol{z}, \boldsymbol{x}, \boldsymbol{\sigma}, \boldsymbol{\zeta}) e(T - t - h) \\ &+ (\hat{\boldsymbol{\lambda}}^m)^\top \tilde{\boldsymbol{f}}(t + h, \boldsymbol{z}, \boldsymbol{x}, \boldsymbol{\sigma}, \boldsymbol{\zeta}) e(T - t - h), \end{aligned} \tag{4.2b}$$

and $e(\cdot)$ is the unit step function.

To continue, we set

$$\boldsymbol{z}(t) = \boldsymbol{0}, \quad \text{for all} \quad t \in [T - h, T]. \tag{4.3}$$

Let $\boldsymbol{\lambda}^m(\cdot \mid \boldsymbol{\sigma}, \boldsymbol{\zeta})$ denote the solution of the co-state system (4.1) corresponding to $(\boldsymbol{\sigma}, \boldsymbol{\zeta}) \in \Theta \times Z$. It is obtained by solving the co-state system (4.1) backward in time from $t = T$ to $t = 0$.

**Theorem 4.1.** Consider Problem (P1). For each $m = 0, 1, \ldots, N$, the gradients of the function $g_m$ with respect to $\boldsymbol{\sigma}$ and $\boldsymbol{\zeta}$ are given by

$$\frac{\partial g_m(\boldsymbol{\sigma},\boldsymbol{\zeta})}{\partial \boldsymbol{\sigma}} = \sum_{k=1}^{M+1} \int_{\tau_{k-1}}^{\tau_k} \frac{\partial H_m}{\partial \boldsymbol{\sigma}} \mathrm{d}t, \tag{4.4a}$$

$$\frac{\partial g_m(\boldsymbol{\sigma},\boldsymbol{\zeta})}{\partial \boldsymbol{\zeta}} = \sum_{k=1}^{M+1} \int_{\tau_{k-1}}^{\tau_k} \frac{\partial H_m}{\partial \boldsymbol{\zeta}} \mathrm{d}t. \tag{4.4b}$$

Proof. The derivation of the gradients of the function $g_m$ with respect to $\boldsymbol{\sigma}$ and $\boldsymbol{\zeta}$ are similar. Thus, only the derivation of the gradient of the function $g_m$ with respect to $\boldsymbol{\zeta}$ is given below. Note that the functions $H_m$ and $\boldsymbol{\lambda}^m$ may have discontinuities at the characteristic time points $\tau_i$, $i = 1, \ldots, M$. Let $\boldsymbol{\zeta} \in Z$ be an arbitrary but fixed parameter vector and let $\boldsymbol{\rho}$ be any perturbation about $\boldsymbol{\zeta}$. Define

$$\boldsymbol{\zeta}(\epsilon) = \boldsymbol{\zeta} + \epsilon\boldsymbol{\rho}, \tag{4.5}$$

where $\epsilon > 0$ is an arbitrarily small real number. For brevity, let $\boldsymbol{x}(\cdot)$ and $\boldsymbol{x}(\cdot;\epsilon)$ denote, respectively, the solutions of system (2.5) with initial conditions (2.2) corresponding to $\boldsymbol{\zeta}$ and $\boldsymbol{\zeta}(\epsilon)$. Let $\boldsymbol{y}(\cdot)$, $\boldsymbol{z}(\cdot)$, $\boldsymbol{y}(\cdot;\epsilon)$, $\boldsymbol{z}(\cdot;\epsilon)$ be as defined according to (4.1e)-(4.1f). From (2.5), we have

$$\boldsymbol{x}(t) = \boldsymbol{x}^0 + \int_0^t \tilde{\boldsymbol{f}}(s, \boldsymbol{x}(s), \boldsymbol{y}(s), \boldsymbol{\sigma}, \boldsymbol{\zeta})\mathrm{d}s \tag{4.6}$$

and

$$\boldsymbol{x}(t;\epsilon) = \boldsymbol{x}^0 + \int_0^t \tilde{\boldsymbol{f}}(s, \boldsymbol{x}(s;\epsilon), \boldsymbol{y}(s;\epsilon), \boldsymbol{\sigma}(\epsilon), \boldsymbol{\zeta}(\epsilon))\mathrm{d}s. \tag{4.7}$$

Thus,

$$\begin{aligned}\Delta\boldsymbol{x}(t) &= \left.\frac{\mathrm{d}\boldsymbol{x}(t;\epsilon)}{\mathrm{d}\epsilon}\right|_{\epsilon=0} \\ &= \int_0^t \left[\frac{\partial \tilde{\boldsymbol{f}}(s, \boldsymbol{x}(s), \boldsymbol{y}(s), \boldsymbol{\sigma}, \boldsymbol{\zeta})}{\partial \boldsymbol{x}}\Delta\boldsymbol{x}(s) + \frac{\partial \tilde{\boldsymbol{f}}(s, \boldsymbol{x}(s), \boldsymbol{y}(s), \boldsymbol{\sigma}, \boldsymbol{\zeta})}{\partial \boldsymbol{y}}\Delta\boldsymbol{y}(s)\right. \\ &\quad \left. + \frac{\partial \tilde{\boldsymbol{f}}(s, \boldsymbol{x}(s), \boldsymbol{y}(s), \boldsymbol{\sigma}, \boldsymbol{\zeta})}{\partial \boldsymbol{\zeta}}\boldsymbol{\rho}\right] \mathrm{d}s. \end{aligned} \tag{4.8}$$

Clearly, $\Delta\boldsymbol{x}(t)$ satisfies

$$\begin{aligned}\frac{\mathrm{d}(\Delta\boldsymbol{x}(t))}{\mathrm{d}t} =& \frac{\partial \tilde{\boldsymbol{f}}(t, \boldsymbol{x}(t), \boldsymbol{y}(t), \boldsymbol{\sigma}, \boldsymbol{\zeta})}{\partial \boldsymbol{x}}\Delta\boldsymbol{x}(t) + \frac{\partial \tilde{\boldsymbol{f}}(t, \boldsymbol{x}(t), \boldsymbol{y}(t), \boldsymbol{\sigma}, \boldsymbol{\zeta})}{\partial \boldsymbol{y}}\Delta\boldsymbol{y}(t) \\ &+ \frac{\partial \tilde{\boldsymbol{f}}(t, \boldsymbol{x}(t), \boldsymbol{y}(t), \boldsymbol{\sigma}, \boldsymbol{\zeta})}{\partial \boldsymbol{\zeta}}\boldsymbol{\rho}, \end{aligned} \tag{4.9a}$$

$$\Delta \boldsymbol{x}(t) = 0, \quad \text{for} \quad t \leq 0. \tag{4.9b}$$

Now, by (2.7), we have

$$\begin{aligned} g_m(\boldsymbol{\sigma}(\epsilon), \boldsymbol{\zeta}(\epsilon)) =& \Phi_m(\boldsymbol{x}(\tau_1;\epsilon), \ldots, \boldsymbol{x}(\tau_{M+1};\epsilon)) \\ &+ \sum_{k=1}^{M+1} \int_{\tau_{k-1}}^{\tau_k} \tilde{\mathcal{L}}_m(t, \boldsymbol{x}(t;\epsilon), \boldsymbol{y}(t;\epsilon), \boldsymbol{\sigma}(\epsilon), \boldsymbol{\zeta}(\epsilon)) \mathrm{d}t. \end{aligned} \tag{4.10}$$

Define

$$\bar{\mathcal{L}}_m = \tilde{\mathcal{L}}_m(t, \boldsymbol{x}(t), \boldsymbol{y}(t), \boldsymbol{\sigma}, \boldsymbol{\zeta}), \quad \hat{\mathcal{L}}_m = \tilde{\mathcal{L}}_m(t+h, \boldsymbol{z}(t), \boldsymbol{x}(t), \boldsymbol{\sigma}, \boldsymbol{\zeta}), \tag{4.11}$$

$$\bar{\boldsymbol{f}} = \tilde{\boldsymbol{f}}(t, \boldsymbol{x}(t), \boldsymbol{y}(t), \boldsymbol{\sigma}, \boldsymbol{\zeta}), \quad \hat{\boldsymbol{f}} = \tilde{\boldsymbol{f}}(t+h, \boldsymbol{z}(t), \boldsymbol{x}(t), \boldsymbol{\sigma}, \boldsymbol{\zeta}), \tag{4.12}$$

$$\bar{H}_m = H_m(t, \boldsymbol{x}(t), \boldsymbol{y}(t), \boldsymbol{\sigma}, \boldsymbol{\zeta}, \boldsymbol{\lambda}^m(t)), \quad \hat{H}_m = \hat{H}_m(t, \boldsymbol{z}(t), \boldsymbol{x}(t), \boldsymbol{\sigma}, \boldsymbol{\zeta}, \hat{\boldsymbol{\lambda}}^m(t)). \tag{4.13}$$

Differentiating (4.10) with respect to $\epsilon$ and then letting $\epsilon = 0$, it follows from using (4.11)-(4.13) that

$$\begin{aligned} \Delta g_m(\boldsymbol{\sigma}, \boldsymbol{\zeta}) = & \left. \frac{\mathrm{d} g_m(\boldsymbol{\sigma}(\epsilon), \boldsymbol{\zeta}(\epsilon))}{\mathrm{d}\epsilon} \right|_{\epsilon=0} = \frac{\partial g_m(\boldsymbol{\sigma}, \boldsymbol{\zeta})}{\partial \boldsymbol{\zeta}} \boldsymbol{\rho} \\ = & \sum_{l=1}^{M+1} \frac{\partial \Phi_m(\boldsymbol{x}(\tau_1), \ldots, \boldsymbol{x}(\tau_{M+1}))}{\partial \boldsymbol{x}(\tau_l)} \Delta \boldsymbol{x}(\tau_l) + \sum_{k=1}^{M+1} \int_{\tau_{k-1}}^{\tau_k} \left[ \frac{\partial \bar{\mathcal{L}}_m}{\partial \boldsymbol{x}} \Delta \boldsymbol{x}(t) \right. \\ & \left. + \frac{\partial \bar{\mathcal{L}}_m}{\partial \boldsymbol{y}} \Delta \boldsymbol{y}(t) + \frac{\partial \bar{\mathcal{L}}_m}{\partial \boldsymbol{\zeta}} \boldsymbol{\rho} \right] \mathrm{d}t. \end{aligned} \tag{4.14}$$

Now, we note that

$$\int_{\tau_{k-1}}^{\tau_k} \left[ \frac{\partial \bar{\mathcal{L}}_m}{\partial \boldsymbol{y}} \Delta \boldsymbol{y}(t) \right] \mathrm{d}t = \int_{\tau_{k-1}}^{\tau_k} \left[ e(\tau_k - t - h) \frac{\partial \hat{\mathcal{L}}_m}{\partial \boldsymbol{x}} \Delta \boldsymbol{x}(t) \right] \mathrm{d}t. \tag{4.15}$$

Combining (4.14) and (4.15) gives

$$\begin{aligned} \Delta g_m(\boldsymbol{\sigma}, \boldsymbol{\zeta}) = & \sum_{l=1}^{M+1} \frac{\partial \Phi_m(\boldsymbol{x}(\tau_1), \ldots, \boldsymbol{x}(\tau_{M+1}))}{\partial \boldsymbol{x}(\tau_l)} \Delta \boldsymbol{x}(\tau_l) + \sum_{k=1}^{M+1} \int_{\tau_{k-1}}^{\tau_k} \left[ \frac{\partial \bar{H}_m}{\partial \boldsymbol{x}} \Delta \boldsymbol{x}(t) \right. \\ & - (\boldsymbol{\lambda}^m)^\top \frac{\partial \bar{\boldsymbol{f}}}{\partial \boldsymbol{x}} \Delta \boldsymbol{x}(t) + \frac{\partial \hat{H}_m}{\partial \boldsymbol{x}} \Delta \boldsymbol{x}(t) - (\hat{\boldsymbol{\lambda}}^m)^\top \frac{\partial \hat{\boldsymbol{f}}}{\partial \boldsymbol{x}} \Delta \boldsymbol{x}(t) e(\tau_k - t - h) \\ & \left. + \frac{\partial \bar{H}_m}{\partial \boldsymbol{\zeta}} \boldsymbol{\rho} - (\boldsymbol{\lambda}^m)^\top \frac{\partial \bar{\boldsymbol{f}}}{\partial \boldsymbol{\zeta}} \boldsymbol{\rho} \right] \mathrm{d}t. \end{aligned} \tag{4.16}$$

From (4.3), (4.1d) and (4.9b), we have

$$\int_{\tau_{k-1}}^{\tau_k}\left[(\hat{\boldsymbol{\lambda}}^m)^\top\frac{\partial\hat{\boldsymbol{f}}}{\partial\boldsymbol{x}}\Delta\boldsymbol{x}(t)e(\tau_k-t-h)\right]\mathrm{d}t=\int_{\tau_{k-1}}^{\tau_k}(\boldsymbol{\lambda}^m)^\top\frac{\partial\bar{\boldsymbol{f}}}{\partial\boldsymbol{y}}\Delta\boldsymbol{y}(t)\mathrm{d}t. \tag{4.17}$$

Thus, it follows from (4.16), (4.17) and (4.9a) that

$$\begin{aligned}\Delta g_m(\boldsymbol{\sigma},\boldsymbol{\zeta})=&\sum_{l=1}^{M+1}\frac{\partial\Phi_m(\boldsymbol{x}(\tau_1),\ldots,\boldsymbol{x}(\tau_{M+1}))}{\partial\boldsymbol{x}(\tau_l)}\Delta\boldsymbol{x}(\tau_l)+\sum_{k=1}^{M+1}\int_{\tau_{k-1}}^{\tau_k}\left[\frac{\partial\bar{H}_m}{\partial\boldsymbol{x}}\Delta\boldsymbol{x}(t)\right.\\&\left.+\frac{\partial\hat{H}_m}{\partial\boldsymbol{x}}\Delta\boldsymbol{x}(t)+\frac{\partial\bar{H}_m}{\partial\boldsymbol{\zeta}}\boldsymbol{\rho}-(\boldsymbol{\lambda}^m)^\top\frac{\mathrm{d}(\Delta\boldsymbol{x}(t))}{\mathrm{d}t}\right]\mathrm{d}t.\end{aligned} \tag{4.18}$$

Applying integration by parts to the last term of the right hand side of the above equation yields

$$\int_{\tau_{k-1}}^{\tau_k}\left[(\boldsymbol{\lambda}^m)^\top\frac{\mathrm{d}(\Delta\boldsymbol{x}(t))}{\mathrm{d}t}\right]\mathrm{d}t=(\boldsymbol{\lambda}^m)^\top\Delta\boldsymbol{x}(t)\big|_{\tau_{k-1}^+}^{\tau_k^-}-\int_{\tau_{k-1}}^{\tau_k}\left[\Delta\boldsymbol{x}(t)\left(\frac{\mathrm{d}\boldsymbol{\lambda}^m}{\mathrm{d}t}\right)^\top\right]\mathrm{d}t. \tag{4.19}$$

From (4.19), it follows from (4.18) that

$$\begin{aligned}\Delta g_m(\boldsymbol{\sigma},\boldsymbol{\zeta})=&\sum_{l=1}^{M+1}\frac{\partial\Phi_m(\boldsymbol{x}(\tau_1),\ldots,\boldsymbol{x}(\tau_{M+1}))}{\partial\boldsymbol{x}(\tau_l)}\Delta\boldsymbol{x}(\tau_l)-\sum_{k=1}^{M+1}(\boldsymbol{\lambda}^m)^\top\Delta\boldsymbol{x}(t)\big|_{\tau_{k-1}^+}^{\tau_k^-}\\&+\sum_{k=1}^{M+1}\int_{\tau_{k-1}}^{\tau_k}\left[\left(\frac{\partial\bar{H}_m}{\partial\boldsymbol{x}}+\frac{\partial\hat{H}_m}{\partial\boldsymbol{x}}+\left(\frac{\mathrm{d}\boldsymbol{\lambda}^m}{\mathrm{d}t}\right)^\top\right)\Delta\boldsymbol{x}(t)+\frac{\partial\bar{H}_m}{\partial\boldsymbol{\zeta}}\boldsymbol{\rho}\right]\mathrm{d}t.\end{aligned} \tag{4.20}$$

Since the state and its gradients with respect to $\boldsymbol{\zeta}$ are continuous in $t$ on $[0,T]$, i.e. $\boldsymbol{x}(\tau_k^-)=\boldsymbol{x}(\tau_k^+)$ and $\Delta\boldsymbol{x}(\tau_k^-)=\Delta\boldsymbol{x}(\tau_k^+)$, for $k=1,\ldots,M$, we obtain

$$\begin{aligned}\sum_{k=1}^{M+1}(\boldsymbol{\lambda}^m)^\top\Delta\boldsymbol{x}(t)\big|_{\tau_{k-1}^+}^{\tau_k^-}=&[\boldsymbol{\lambda}^m(\tau_{M+1}^-)^\top\Delta\boldsymbol{x}(\tau_{M+1}^-)-\boldsymbol{\lambda}^m(\tau_M^+)^\top\Delta\boldsymbol{x}(\tau_M^+)]+[\boldsymbol{\lambda}^m(\tau_M^-)^\top\Delta\boldsymbol{x}(\tau_M^-)\\&-\boldsymbol{\lambda}^m(\tau_{M-1}^+)^\top\Delta\boldsymbol{x}(\tau_{M-1}^+)]+\ldots+[\boldsymbol{\lambda}^m(\tau_1^-)^\top\Delta\boldsymbol{x}(\tau_1^-)\\&-\boldsymbol{\lambda}^m(\tau_0^+)^\top\Delta\boldsymbol{x}(\tau_0^+)]\\=&\sum_{k=1}^{M}[\boldsymbol{\lambda}^m(\tau_k^-)^\top-\boldsymbol{\lambda}^m(\tau_k^+)^\top]\Delta\boldsymbol{x}(\tau_k)-\boldsymbol{\lambda}^m(\tau_0^+)^\top\Delta\boldsymbol{x}(\tau_0^+)\\&+\boldsymbol{\lambda}^m(\tau_{M+1}^-)^\top\Delta\boldsymbol{x}(\tau_{M+1}^-)\\=&\sum_{k=1}^{M}[\boldsymbol{\lambda}^m(\tau_k^-)^\top-\boldsymbol{\lambda}^m(\tau_k^+)^\top]\Delta\boldsymbol{x}(\tau_k)+\boldsymbol{\lambda}^m(T)^\top\Delta\boldsymbol{x}(T).\end{aligned} \tag{4.21}$$

Substituting (4.21) into (4.20) and noting that $\boldsymbol{x}(\tau_0^+) = \boldsymbol{x}(0) = \boldsymbol{x}^0$, which is a fixed vector in $\mathbb{R}^n$, it follows that

$$\begin{aligned}\Delta g_m(\boldsymbol{\sigma},\boldsymbol{\zeta}) =& \sum_{k=1}^{M}\left[\frac{\partial\Phi_m(\boldsymbol{x}(\tau_1),\ldots,\boldsymbol{x}(\tau_{M+1}))}{\partial\boldsymbol{x}(\tau_k)} - \boldsymbol{\lambda}^m(\tau_k^-)^\top + \boldsymbol{\lambda}^m(\tau_k^+)^\top\right]\Delta\boldsymbol{x}(\tau_k)\\ &+ \frac{\partial\Phi_m(\boldsymbol{x}(\tau_1),\ldots,\boldsymbol{x}(\tau_{M+1}))}{\partial\boldsymbol{x}(T)}\Delta\boldsymbol{x}(T) - \boldsymbol{\lambda}^m(T)^\top\Delta\boldsymbol{x}(T)\\ &+ \sum_{k=1}^{M+1}\int_{\tau_{k-1}}^{\tau_k}\left[\left(\frac{\partial\bar{H}_m}{\partial\boldsymbol{x}} + \frac{\partial\hat{H}_m}{\partial\boldsymbol{x}} + \left(\frac{\mathrm{d}\boldsymbol{\lambda}^m}{\mathrm{d}t}\right)^\top\right)\Delta\boldsymbol{x}(t) + \frac{\partial\bar{H}_m}{\partial\boldsymbol{\zeta}}\boldsymbol{\rho}\right]\mathrm{d}t. \quad (4.22)\end{aligned}$$

By virtue of the definition of the co-state system corresponding to $g_m$ given in (4.1a) with the jump conditions (4.1b) and terminal condition (4.1c), we obtain

$$\Delta g_m(\boldsymbol{\sigma},\boldsymbol{\zeta}) = \frac{\partial g_m(\boldsymbol{\sigma},\boldsymbol{\zeta})}{\partial\boldsymbol{\zeta}}\boldsymbol{\rho} = \sum_{k=1}^{M+1}\int_{\tau_{k-1}}^{\tau_k}\left(\frac{\partial\bar{H}_m}{\partial\boldsymbol{\zeta}}\boldsymbol{\rho}\right)\mathrm{d}t. \quad (4.23)$$

Since $\boldsymbol{\rho}$ is arbitrary, we obtain the gradient formula (4.4b). The gradient formula (4.4a) can be derived similarly. This completes the proof.

**Remark 4.1.** In view of the co-state systems (4.1), we observe that their trajectories are discontinuous at the characteristic time points. The sizes of the jumps are determined by the jump point conditions given by (4.1b).

# 5 Numerical simulation for industrial application

To illustrate the applicability of the solution method discussed in the preceding sections, we consider an industrial application involving the purification process of zinc sulphate electrolyte.

In the purification process of zinc hydrometallurgical production, the presence of metallic impurities in zinc sulphate electrolyte is a major concern for the zinc electrolysis process. These impurities, which are typically cobalt and cadmium ions in the electrolyte tank, can decrease the efficiency of the current flow and cause the dissolution of the cathode during the electrolysis. Consequently, the purity grade of deposited zinc is reduced. Therefore, adequate purification of the zinc sulphate electrolyte before electrolysis is essential. This purification is carried out by deposition, which involves the removal of metallic impurities especially cobalt and cadmium ions from the zinc sulphate electrolyte by deposition on zinc.

In the practical industrial process of a zinc production factory in China, the zinc powder is added continuously to the zinc sulphate electrolyte at the inlet of the purification reaction tank to remove the metallic impurities ions. Generally, only the concentrations of cobalt and cadmium ions at the inlet and outlet of the reaction tank are measured at each hour. Meanwhile it usually takes two hours for zinc sulphate electrolyte to flow from the inlet to outlet of the reaction tank. In other words, the reaction time for the purification process is two hours. Due to the complicated reaction environment and uncertainties, the zinc powder is added much more excessively than required so as to guarantee the quality of the production. Thus, there is much waste in zinc powder. Clearly, this is not desirable in practice.

Based on the chemical kinetics (see [19, 20]), the chemical reaction equations which are relevant to zinc-cobalt and zinc-cadmium deposition reaction can be described by the following equations.

$$V\frac{\mathrm{d}x_1(t)}{\mathrm{d}t} = Qx_{10} - Qx_1(t-2) - \alpha u(t)x_1(t-2) + cx_2(t-2), \tag{5.1}$$

$$V\frac{\mathrm{d}x_2(t)}{\mathrm{d}t} = Qx_{20} - Qx_2(t-2) - \beta v(t)x_2(t-2) + dx_1(t-2). \tag{5.2}$$

These equations are obtained through the knowledge of continuous stirred tank reactor (CSTR) and replacement reaction of metallic ions. These chemical reaction equations are similar because the reaction properties of zinc-cobalt and zinc-cadmium are similar. In the above equations, $x_1(t)$ and $x_2(t)$ represent, respectively, the concentrations of cobalt and cadmium ions in the reaction tank, $x_{10}$ and $x_{20}$ represent, respectively, the concentrations of cobalt and cadmium ions at the inlet of the reaction tank, $Q$ is the flux of solution, $V$ is the volume of the reaction tank. $\alpha$ and $\beta$ are, respectively, the chemical reaction coefficients for cobalt and cadmium ions. $u$ and $v$ are the zinc powder reaction surface areas for two metallic impurities ions, they are control variables and can be converted into the weights of zinc powder because the relationship between reaction surface area and weight is linear as suggested from experimental validation. $c$ and $d$ are coupling coefficients. Here, $x_{10}, x_{20}, Q$ and $V$ are known parameters. Let $\hat{x}_1(i)$ and $\hat{x}_2(i)$ denote, respectively, the observed data of concentrations of cobalt and cadmium ions at the measurement time points $i = 1, \ldots, N$. Our first aim is to choose the optimal system parameters $\alpha, \beta, c, d$, so that the trajectory "best" fits the observed data at these measurement time points. That is, the following cost function

$$J_1 = \sum_{i=1}^{N}[(x_1(i) - \hat{x}_1(i))^2 + (x_2(i) - \hat{x}_2(i))^2] \tag{5.3}$$

is minimized with respect to the parameters $\alpha, \beta, c, d$, where $x_1(i)$ and $x_2(i)$ are, respectively, the concentrations of cobalt and cadmium ions at the time points $i = 1, \ldots, N$ calculated from the solution of the system of time delayed differential equations corresponding to the choice of parameters $\alpha, \beta, c, d$. The states before zero time point are obtained through interpolation of the measured data at the measurement time points prior to the zero time point. The cubic basis spline interpolation method (see [17]) is adopted to construct the fitting curve which passes through all the measurement data. Now, the solution method proposed in Section 4 is used to minimize the cost function (5.3) with respect to the system parameters $\alpha, \beta, c, d$. The optimal solutions obtained are $\alpha = 5.464 \times 10^{-4}$, $\beta = 3.664 \times 10^{-4}$, $c = 9.54$, $d = 1.415 \times 10^3$. Let $\bar{x}_i$, $i = 1, 2$, be the concentrations of cobalt and cadmium ions corresponding to these optimal parameters. They are shown as dashed lines in Figures 1-2. However, the zinc powder reaction surface areas used are rather excessive than required. We thus move to minimize the zinc powder reaction surface areas subject to the condition that the possible deviations of $x_i(t)$, $i = 1, 2$, which are the solutions of the system of time delayed differential equations corresponding to the choice of the controls $u$ and $v$, from $\bar{x}_i(t)$, $i = 1, 2$, are within an acceptable limit.

This can be posed as a time delayed optimal control problem in the form of Problem (P2), where

$$J_2 = \int_0^T (u(t)^2 + v(t)^2)\mathrm{d}t \tag{5.4}$$

is minimized, subject to the continuous state inequality constraints

$$(x_i(t) - \bar{x}_i(t))^2 \leq e, \quad t \in [0, T], \quad i = 1, 2. \tag{5.5}$$

Here, $e > 0$ indicates the acceptable limit.

As shown in Section 3, this optimal control problem is approximated by a sequence of Problems (P2($\varepsilon, \gamma$)). The optimal concentrations of cobalt and cadmium ions obtained by using the algorithm given in Section 4 are shown as solid lines in Figures 1-2. We see that these optimal concentrations remain on track to the measured data. Several measured data deviate rather far away from the calculated concentration. This is expected, as the measurements are carried out manually by using some special chemical instruments. Inevitably, it will lead to some measurement inaccuracy. However, the trajectory obtained does reflect the basic characteristic of the practical scenarios under study.

The respective reaction surface areas of zinc powder for cobalt and cadmium ion, which are the control functions, are shown in Figures 3-4. The total reaction surface area of zinc powder is shown in Figure 5. We see that the average reaction surface area is much lower than the one used in current practice, shown as dashed line. As mentioned earlier, the amount of zinc powder actually added is expressed as a whole quantity and the reaction surface area can be converted to the weight according to the linear relationship between the weight and reaction surface area of zinc powder. Our proposed method has found the optimal consumption of zinc powder at each hour effectively.

# 6 Conclusion

This paper considered a class of optimal control problems involving time delayed systems and subject to inequality constraints in canonical form, where the cost and con-

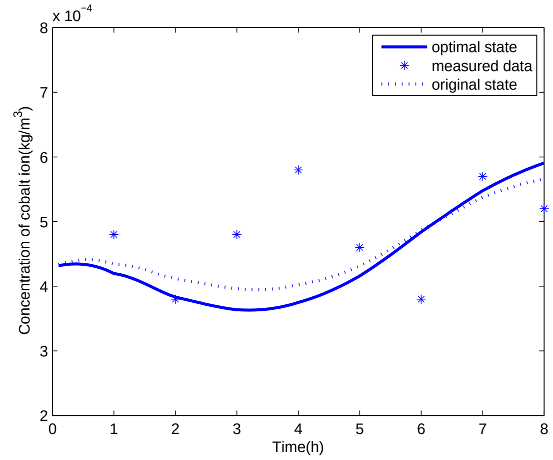


Figure 1: Concentration of cobalt ion.

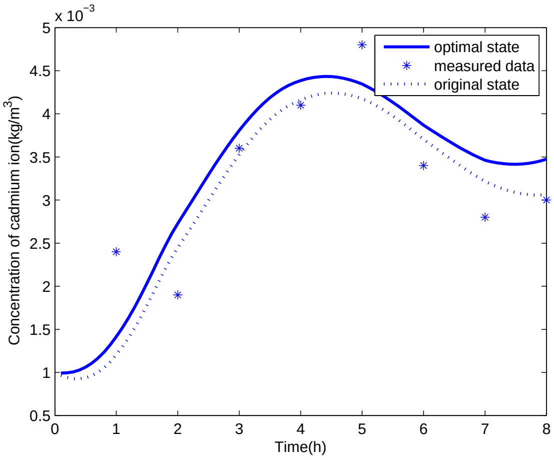


Figure 2: Concentration of cadmium ion.

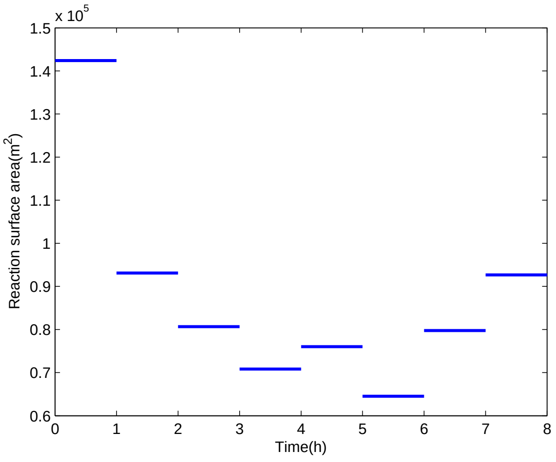


Figure 3: Reaction surface area of zinc powder for cobalt ion.

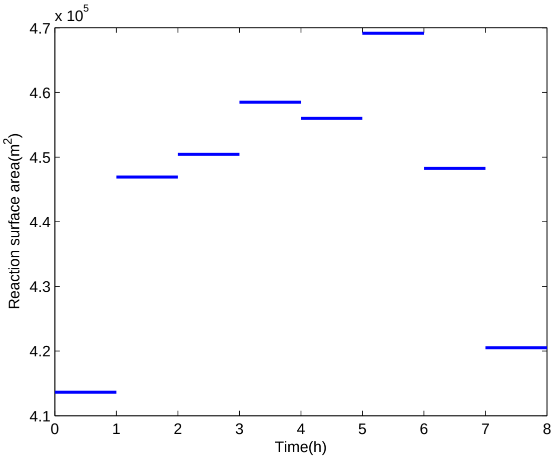


Figure 4: Reaction surface area of zinc powder for cadmium ion.

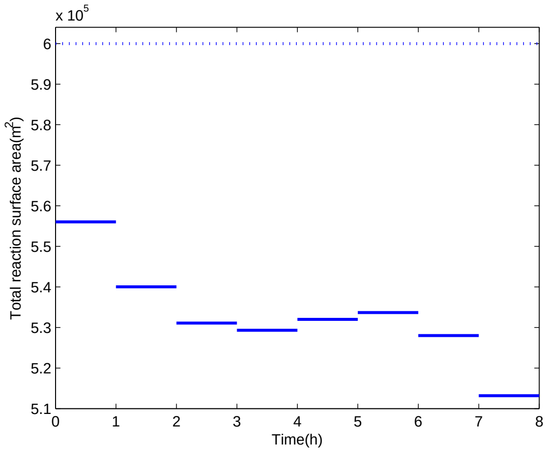


Figure 5: Total reaction surface area of zinc powder.

straint functions are depending on multiple characteristic time points. For a time delayed optimal control problem subject to continuous state inequality constraints, it was shown that this optimal control problem can be approximated by a sequence of approximate optimal control problems subject to inequality constraints in canonical form. Gradient formulas for the cost and constraint functions were derived. On this basis, the time delayed optimal control problem subject to inequality constraints in canonical form can be solved as a nonlinear optimization problem by using any gradient-based method. The obtained computational method was applied to the study of purification process of zinc sulphate electrolyte. The results obtained were highly satisfactory, showing the effectiveness of the method proposed.

# Acknowledgment

This work was supported by the Australian Research Council and National Natural Science Foundation of China under Grant No. 60634020, 60704003 and 60874069. We would like to thank the anonymous referees for their constructive comments and suggestions.

# References

[1] B. S. Chen, S. S. Wang, H. C. Lu, Stabilization of time-delay systems containing saturating actuators, International Journal of Control, 47(3)(1988), 867-881.

[2] D. H. Chyung, E. B. Lee, Linear optimal systems with time-delay, SIAM Journal on Control, 4(3)(1966), 548-575.

[3] T. N. Lee, U. L. Radovic, General decentralized stabilization of large-scale linear continuous and discrete time-delay systems, International Journal of Control, 46(3)(1987), 2127-2140.

[4] J. N. Yang, T. T. Soong, Recent advances in active control of civil engineering structures, Probabilistic Engineering Mechanics, 3(4)(1988), 179-188.

[5] G. P. Cai, J. Z. Huang, S. X. Yang, An Optimal control method for linear systems with time Delay, Computers and Structures, 81(2003), 1539-1546.

[6] K. H. Wong, D. J. Clements, K. L. Teo, Optimal control computation for nonlinear time-lag systems, Journal of Optimization Theory and Applications, 47(1)(1985), 91-107.

[7] K. L. Teo, K. H. Wong, D. J. Clements, Optimal control computation for linear time-lag systems with linear terminal constraints, Journal of Optimization Theory and Applications, 44(3)(1984), 509-526.

[8] S. Dadebo, R. Luus, Optimal control of time-delay systems by dynamic programming, Optimal Control Applications and Methods, 13(1992), 29-41.

[9] R. Luus, X. Zhang, F. Hartig, F. J. Keil, Use of piecewise linear continuous optimal control for time-delay systems, Industrial & Engineering Chemistry Research, 34(1995), 4136-4139.

[10] K. L. Teo, C. J. Goh, A computational method for combined optimal parameter selection and optimal control problems with general constraints, Journal of the Australian Mathematical Society, Series B, 30(1989), 350-364.

[11] R. B. Martin, Optimal control of cancer chemotherapy, Automatica, 28(6)(1992), 1113-1123.

[12] R. B. Martin, K. L. Teo, *Optimal control of drug administration in cancer chemotherapy*. Singapore: World Scientific, 1994.

[13] R. C. Loxton, K. L. Teo, V. Rehbock, Optimal control problems with multiple characteristic time points in the objective and constraints, Automatica, 44(2008), 2923-2929.

[14] R. Li, Z. G. Feng, K. L. Teo, G. R. Duan, Optimal piecewise state feedback control for impulsive switched system, Mathematical and Computer Modelling, 48(2008), 468-479.

[15] K. Kaji, K. H. Wong, Nonlinearly constrained time-delayed optimal control problems, Journal of Optimization Theory and Applications, 82(2)(1994), 295-313.

[16] K. L. Teo, C. J. Goh, K. H. Wong. *A Unified computational approach to optimal control problems*. Longman Scientific and Technical, New York, 1991.

[17] V. Rehbock, K. L. Teo, L. S. Jennings, Suboptimal feedback control for a class of nonlinear systems using spline interpolation, Discrete and Continuous Dynamical Systems, 1(2)(1995), 223-236.

[18] K. Schittkowski, NLPQLP: A Fortran Implementation of a Sequential Quadratic Programming Algorithm with Distributed and Non-monotone Line Search - User's Guide Version 2.0, University of Bayreuth, 2004.

[19] A. M. Polcaro, S. Palmas, S. Dernini, Kinetics of cobalt cementation on zinc powder, Industrial & Engineering Chemistry Research, 34(1995), 3090-3095.

[20] W. L. Choo, M. I. Jeffrey, S. G. Robertson, Analysis of leaching and cementation reaction kinetics: correcting for volume changes in laboratory studies, Hydrometallurgy, 82(2006), 110-116.