A class of optimal state-delay control problems

Qinqin Chai[a,b,*], Ryan Loxton[b], Kok Lay Teo[b], Chunhua Yang[a]

[a] *School of Information Science & Engineering, Central South University, Changsha, P. R. China*
[b] *Department of Mathematics & Statistics, Curtin University, Perth, Australia*

Abstract

We consider a general nonlinear time-delay system with state-delays as control variables. The problem of determining optimal values for the state-delays to minimize overall system cost is a non-standard optimal control problem—called an optimal state-delay control problem—that cannot be solved using existing techniques. We show that this optimal control problem can be formulated as a nonlinear programming problem in which the cost function is an implicit function of the decision variables. We then develop an efficient numerical method for determining the cost function's gradient. This method, which involves integrating an impulsive dynamic system backwards in time, can be combined with any standard gradient-based optimization method to solve the optimal state-delay control problem effectively. We conclude the paper by discussing applications of our approach to parameter identification and delayed feedback control.

Keywords: time-delay, optimal control, nonlinear optimization, parameter identification, delayed feedback control

---

[*] Corresponding author
Email addresses: kppqing@163.com (Qinqin Chai), r.loxton@curtin.edu.au (Ryan Loxton), k.l.teo@curtin.edu.au (Kok Lay Teo), ychh@csu.edu.cn (Chunhua Yang)

1. Introduction

Time-delay systems arise in many real-world applications—e.g. evaporation and purification processes [1, 2], aerospace models [3], and human immune response [4]. Over the past two decades, various optimal control methods have been developed for time-delay systems. Well-known tools include the necessary conditions for optimality [5, 6] and numerical methods based on the control parameterization technique [7, 8]. These existing optimal control methods are restricted to time-delay systems in which the delays are fixed and known. In this paper, we consider a new class of optimal control problems in which the delays are not fixed, but are instead control variables to be chosen optimally. Such problems are called *optimal state-delay control problems*.

As an example of an optimal state-delay control problem, consider a system of delay-differential equations with unknown delays. This delay-differential system is a dynamic model for some phenomenon under consideration. The problem is to choose values for the unknown delays (and possibly other model parameters) so that the system output predicted by the model is consistent with experimental data. This so-called *parameter identification problem* can be formulated as an optimal state-delay control problem in which the delays and model parameters are decision variables, and the cost function measures the least-squares error between predicted and observed system output.

Parameter identification for time-delay systems has been an active area of research over the past decade. Existing techniques for parameter identification include interpolation methods [9], genetic algorithms [10], and the delay operator transform method [11]. These techniques are mainly designed for single-delay linear systems. In contrast, the computational approach to be developed in this paper, which is based on formulating and solving the parameter identification problem as an optimal state-delay control problem, can handle systems with nonlinear dynamics and multiple time-delays. This computational approach is motivated by our earlier work in [12], which presents a parameter identification algorithm based on nonlinear programming techniques. This algorithm has two limitations: (i) it is only applicable to systems in which each nonlinear term contains a single delay and no unknown parameters; and (ii) it involves integrating a large number of auxiliary delay-differential systems (one auxiliary system for each unknown delay and model parameter). The new approach to be developed in this paper does not suffer from these limitations. In particular, our new approach only requires the integration of one auxiliary system, regardless of the number of delays and parameters in the underlying dynamic model.

Another important application of optimal state-delay control problems is in delayed feedback control. In delayed feedback control, the system's input function is chosen to be a linear function of the delayed state, as opposed to traditional feedback control in which the input is a function of the current (undelayed) state. Voluntarily introducing delays via delayed feedback control can be beneficial for certain types of systems; see, for example, [13, 14, 15]. The problem of choosing optimal values for the delays in a delayed feedback controller can be formulated as an optimal state-delay control problem.

Our goal in this paper is to develop a unified computational approach for solving optimal state-delay control problems. A key aspect of our work is the derivation of an *auxiliary impulsive system*, which turns out to be the analogue of the costate system in classical optimal control. We derive formulae for the cost function's gradient in terms of the solution of this impulsive system. On this basis, the optimal state-delay control problem can be solved by combining numerical integration and nonlinear programming techniques. This approach has proven very effective for the two specific applications mentioned above—parameter identification and delayed feedback control.

The remainder of the paper is organized as follows. We first formulate the optimal state-delay control problem in Section 2, before introducing the auxiliary impulsive system and deriving gradient formular in Section 3. Section 4 is devoted to parameter identification problems, and Section 5 is devoted to delayed feedback control. We make some concluding remarks in Section 6.

2. Problem formulation

Consider the following nonlinear time-delay system:

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{f}(\boldsymbol{x}(t), \boldsymbol{x}(t-\tau_1), \ldots, \boldsymbol{x}(t-\tau_m), \boldsymbol{\zeta}), \quad t \in [0, T], \tag{1}$$

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

where $T > 0$ is a given *terminal time*; $\boldsymbol{x}(t) = [x_1(t), \ldots, x_n(t)]^\top \in \mathbb{R}^n$ is the *state vector*; τ_i, $i = 1, \ldots, m$ are *state-delays*; $\boldsymbol{\zeta} = [\zeta_1, \ldots, \zeta_r]^\top \in \mathbb{R}^r$ is a vector of *system parameters*; and $\boldsymbol{f} : \mathbb{R}^{(m+1)n} \times \mathbb{R}^r \to \mathbb{R}^n$ and $\boldsymbol{\phi} : \mathbb{R} \times \mathbb{R}^r \to \mathbb{R}^n$ are given functions.

System (1)-(2) is controlled via the state-delays and system parameters—these must be chosen optimally so that the system behaves in the best possible manner. We impose the following bound constraints:

$$a_i \leq \tau_i \leq b_i, \quad i = 1, \ldots, m, \tag{3}$$

and

$$c_j \leq \zeta_j \leq d_j, \quad j = 1, \ldots, r, \tag{4}$$

where a_i and b_i are given constants such that $0 \leq a_i < b_i$, and c_j and d_j are given constants such that $c_j < d_j$.

Any vector $\boldsymbol{\tau} = [\tau_1, \ldots, \tau_m]^\top \in \mathbb{R}^m$ satisfying (3) is called an *admissible state-delay vector*. Let $\mathcal{T}$ denote the set of all such admissible state-delay vectors.

Any vector $\boldsymbol{\zeta} = [\zeta_1, \ldots, \zeta_r]^\top \in \mathbb{R}^r$ satisfying (4) is called an *admissible parameter vector*. Let $\mathcal{Z}$ denote the set of all such admissible parameter vectors.

Any combined pair $(\boldsymbol{\tau}, \boldsymbol{\zeta}) \in \mathcal{T} \times \mathcal{Z}$ is called an *admissible control pair* for system (1)-(2).

We assume that the following conditions are satisfied.

Assumption 1. *The given function* $\boldsymbol{f}$ *is continuously differentiable, and* $\boldsymbol{\phi}$ *is twice continuously differentiable.*

Assumption 2. *There exists a real number* $L_1 > 0$ *such that for all* $\boldsymbol{\xi}^i \in \mathbb{R}^n$, $i = 0, \ldots, m$, *and* $\boldsymbol{\omega} \in \mathbb{R}^r$,

$$|\boldsymbol{f}(\boldsymbol{\xi}^0, \boldsymbol{\xi}^1, \ldots, \boldsymbol{\xi}^m, \boldsymbol{\omega})| \leq L_1(1 + |\boldsymbol{\xi}^0| + |\boldsymbol{\xi}^1| + \cdots + |\boldsymbol{\xi}^m| + |\boldsymbol{\omega}|),$$

where $|\cdot|$ *denotes the Euclidean norm.*

Assumptions 1 and 2 ensure that system (1)-(2) admits a unique solution corresponding to each admissible control pair $(\boldsymbol{\tau}, \boldsymbol{\zeta}) \in \mathcal{T} \times \mathcal{Z}$ [16]. We denote this solution by $\boldsymbol{x}(\cdot|\boldsymbol{\tau}, \boldsymbol{\zeta})$.

Our aim is to find an admissible control pair that minimizes the following cost function:

$$J(\boldsymbol{\tau}, \boldsymbol{\zeta}) = \Phi(\boldsymbol{x}(t_1|\boldsymbol{\tau}, \boldsymbol{\zeta}), \ldots, \boldsymbol{x}(t_p|\boldsymbol{\tau}, \boldsymbol{\zeta}), \boldsymbol{\zeta}), \tag{5}$$

where $\Phi : \mathbb{R}^{pn} \times \mathbb{R}^r \to \mathbb{R}$ is a given function and t_k, $k = 1, \ldots, p$ are given time points satisfying

$$0 < t_1 < \cdots < t_p \leq T.$$

Unlike the standard Mayer cost function commonly used in optimal control (which depends solely on the final state), the cost function (5) depends on the state at a set of intermediate time points t_k, $k = 1, \ldots, p$. These time points are called *characteristic times* in the optimal control literature [2, 17, 18]. As we will see, cost functions with characteristic times arise in parameter identification problems, where the aim is to minimize the discrepancy between predicted and observed system output at a set of sample times.

Our optimal state-delay control problem is defined formally below.

Problem (P). *Choose* $(\boldsymbol{\tau}, \boldsymbol{\zeta}) \in \mathcal{T} \times \mathcal{Z}$ *to minimize the cost function* (5).

3. Gradient computation

Although the optimal control of time-delay systems has been the subject of numerous theoretical and practical investigations [2, 8, 19, 5], most research has focussed on the simple case when the delays are fixed and known. The delays in Problem (P), however, are actually control variables to be determined optimally. Hence, Problem (P) differs considerably from most time-delay optimal control problems considered in the literature.

The aim of this paper is to develop a computational method for solving Problem (P). Our approach is based on the following key observation: Problem (P) can be viewed as a nonlinear optimization problem in which the decision vectors $\boldsymbol{\tau}$ and $\boldsymbol{\zeta}$ influence the cost function J *implicitly* through the governing dynamic system (1)-(2). Thus, if the gradient of J can be computed for each admissible control pair, then Problem (P) can be solved using existing gradient-based optimization methods, such as sequential quadratic programming (see [20, 21]). However, since J is not an explicit function of $\boldsymbol{\tau}$ and $\boldsymbol{\zeta}$, deriving its gradient is not straightforward. The purpose of this section is to develop a numerical algorithm for computing the gradient of J.

3.1. Gradient with respect to state-delays

Define

$$\boldsymbol{\psi}(t|\boldsymbol{\tau},\boldsymbol{\zeta}) = \begin{cases} \frac{\partial \boldsymbol{\phi}(t,\boldsymbol{\zeta})}{\partial t}, & \text{if } t \leq 0, \\ \boldsymbol{f}(\boldsymbol{x}(t|\boldsymbol{\tau},\boldsymbol{\zeta}), \boldsymbol{x}(t-\tau_1|\boldsymbol{\tau},\boldsymbol{\zeta}), \ldots, \boldsymbol{x}(t-\tau_m|\boldsymbol{\tau},\boldsymbol{\zeta}), \boldsymbol{\zeta}), & \text{if } t \in (0,T]. \end{cases}$$

Furthermore, define

$$\frac{\partial \bar{\boldsymbol{f}}(t|\boldsymbol{\tau},\boldsymbol{\zeta})}{\partial \boldsymbol{x}} = \frac{\partial \boldsymbol{f}(\boldsymbol{x}(t|\boldsymbol{\tau},\boldsymbol{\zeta}), \boldsymbol{x}(t-\tau_1|\boldsymbol{\tau},\boldsymbol{\zeta}), \ldots, \boldsymbol{x}(t-\tau_m|\boldsymbol{\tau},\boldsymbol{\zeta}), \boldsymbol{\zeta})}{\partial \boldsymbol{x}},$$
$$\frac{\partial \bar{\boldsymbol{f}}(t|\boldsymbol{\tau},\boldsymbol{\zeta})}{\partial \tilde{\boldsymbol{x}}^i} = \frac{\partial \boldsymbol{f}(\boldsymbol{x}(t|\boldsymbol{\tau},\boldsymbol{\zeta}), \boldsymbol{x}(t-\tau_1|\boldsymbol{\tau},\boldsymbol{\zeta}), \ldots, \boldsymbol{x}(t-\tau_m|\boldsymbol{\tau},\boldsymbol{\zeta}), \boldsymbol{\zeta})}{\partial \tilde{\boldsymbol{x}}^i},$$

where $\frac{\partial}{\partial \tilde{\boldsymbol{x}}^i}$ denotes differentiation with respect to the ith delayed state vector.

Consider the following impulsive dynamic system:

$$\dot{\boldsymbol{\lambda}}(t) = -\left[\frac{\partial \bar{\boldsymbol{f}}(t|\boldsymbol{\tau},\boldsymbol{\zeta})}{\partial \boldsymbol{x}}\right]^\top \boldsymbol{\lambda}(t) - \sum_{l=1}^{m}\left[\frac{\partial \bar{\boldsymbol{f}}(t+\tau_l|\boldsymbol{\tau},\boldsymbol{\zeta})}{\partial \tilde{\boldsymbol{x}}^l}\right]^\top \boldsymbol{\lambda}(t+\tau_l), \tag{6}$$

$$\boldsymbol{\lambda}(t_k^-) = \boldsymbol{\lambda}(t_k^+) + \left[\frac{\partial \Phi(\boldsymbol{x}(t_1|\boldsymbol{\tau},\boldsymbol{\zeta}),\ldots,\boldsymbol{x}(t_p|\boldsymbol{\tau},\boldsymbol{\zeta}),\boldsymbol{\zeta})}{\partial \boldsymbol{x}(t_k)}\right]^\top, \quad k=1,\ldots,p, \tag{7}$$

$$\boldsymbol{\lambda}(t) = \boldsymbol{0}, \quad t \geq t_p. \tag{8}$$

Let $\boldsymbol{\lambda}(\cdot|\boldsymbol{\tau},\boldsymbol{\zeta})$ denote the solution of system (6)-(8) corresponding to the admissible control pair $(\boldsymbol{\tau},\boldsymbol{\zeta}) \in \mathcal{T} \times \mathcal{Z}$.

The following result gives formulae for the partial derivatives of J with respect to the state-delays.

Theorem 1. *For each $(\boldsymbol{\tau},\boldsymbol{\zeta}) \in \mathcal{T} \times \mathcal{Z}$ and i=1,...,m,*

$$\frac{\partial J(\boldsymbol{\tau},\boldsymbol{\zeta})}{\partial \tau_i} = -\int_0^{t_p} \boldsymbol{\lambda}^\top(t|\boldsymbol{\tau},\boldsymbol{\zeta})\frac{\partial \bar{\boldsymbol{f}}(t|\boldsymbol{\tau},\boldsymbol{\zeta})}{\partial \tilde{\boldsymbol{x}}^i}\boldsymbol{\psi}(t-\tau_i|\boldsymbol{\tau},\boldsymbol{\zeta})dt. \tag{9}$$

Proof. Let $\boldsymbol{v} : [0,\infty) \to \mathbb{R}^n$ be an arbitrary function satisfying the following conditions:

(i) $\boldsymbol{v}$ is continuous on the intervals (t_{k-1}, t_k), $k = 1,\ldots,p$, where $t_0 = 0$ by convention;

(ii) $\boldsymbol{v}$ is differentiable almost everywhere;

(iii) $\boldsymbol{v}$ has finite left and right limits at $t = t_k$, $k = 1,\ldots,p$, and a finite right limit at $t = 0$.

Note that any discontinuity of $\boldsymbol{v}$ must lie in the set $\{t_0, t_1, \ldots, t_p\}$.

We may express the cost function J as follows:

$$
\begin{aligned}
J(\boldsymbol{\tau}, \boldsymbol{\zeta}) &= \Phi(\boldsymbol{x}(t_1), \ldots, \boldsymbol{x}(t_p), \boldsymbol{\zeta}) \\
&= \Phi(\boldsymbol{x}(t_1), \ldots, \boldsymbol{x}(t_p), \boldsymbol{\zeta}) \\
&\quad + \int_0^{t_p} \big(\boldsymbol{v}^\top(t)\boldsymbol{f}(\boldsymbol{x}(t), \boldsymbol{x}(t-\tau_1), \ldots, \boldsymbol{x}(t-\tau_m), \boldsymbol{\zeta}) - \boldsymbol{v}^\top(t)\dot{\boldsymbol{x}}(t)\big) dt \\
&= \Phi(\boldsymbol{x}(t_1), \ldots, \boldsymbol{x}(t_p), \boldsymbol{\zeta}) - \sum_{k=1}^{p} \int_{t_{k-1}}^{t_k} \boldsymbol{v}^\top(t)\dot{\boldsymbol{x}}(t) dt \\
&\quad + \sum_{k=1}^{p} \int_{t_{k-1}}^{t_k} \boldsymbol{v}^\top(t)\boldsymbol{f}(\boldsymbol{x}(t), \boldsymbol{x}(t-\tau_1), \ldots, \boldsymbol{x}(t-\tau_m), \boldsymbol{\zeta}) dt,
\end{aligned}
$$

where for simplicity we have omitted the $\boldsymbol{\tau}$ and $\boldsymbol{\zeta}$ arguments in $\boldsymbol{x}(\cdot|\boldsymbol{\tau}, \boldsymbol{\zeta})$. This notation will not cause confusion because $\boldsymbol{\tau}$ and $\boldsymbol{\zeta}$ are assumed to be fixed throughout this proof (in the sequel, we will also omit the $\boldsymbol{\tau}$ and $\boldsymbol{\zeta}$ arguments from $\frac{\partial \bar{\boldsymbol{f}}(t|\boldsymbol{\tau},\boldsymbol{\zeta})}{\partial \boldsymbol{x}}$, $\frac{\partial \bar{\boldsymbol{f}}(t|\boldsymbol{\tau},\boldsymbol{\zeta})}{\partial \tilde{\boldsymbol{x}}^i}$, and $\boldsymbol{\psi}(t|\boldsymbol{\tau}, \boldsymbol{\zeta})$).

Applying integration by parts to the last integral gives

$$
\begin{aligned}
J(\boldsymbol{\tau}, \boldsymbol{\zeta}) &= \Phi(\boldsymbol{x}(t_1), \ldots, \boldsymbol{x}(t_p), \boldsymbol{\zeta}) \\
&\quad + \sum_{k=1}^{p} \int_{t_{k-1}}^{t_k} \boldsymbol{v}^\top(t)\boldsymbol{f}(\boldsymbol{x}(t), \boldsymbol{x}(t-\tau_1), \ldots, \boldsymbol{x}(t-\tau_m), \boldsymbol{\zeta}) dt \\
&\quad - \sum_{k=1}^{p} \Big\{\boldsymbol{v}^\top(t_k^-)\boldsymbol{x}(t_k) - \boldsymbol{v}^\top(t_{k-1}^+)\boldsymbol{x}(t_{k-1})\Big\} + \sum_{k=1}^{p} \int_{t_{k-1}}^{t_k} \dot{\boldsymbol{v}}^\top(t)\boldsymbol{x}(t) dt.
\end{aligned}
\tag{10}
$$

Consider the third term on the right-hand side of (10):

$$\begin{aligned}
&\sum_{k=1}^{p}\Big\{\boldsymbol{v}^\top(t_k^-)\boldsymbol{x}(t_k)-\boldsymbol{v}^\top(t_{k-1}^+)\boldsymbol{x}(t_{k-1})\Big\}\\
&\qquad=\sum_{k=1}^{p}\boldsymbol{v}^\top(t_k^-)\boldsymbol{x}(t_k)-\sum_{k=1}^{p}\boldsymbol{v}^\top(t_{k-1}^+)\boldsymbol{x}(t_{k-1})\\
&\qquad=\sum_{k=1}^{p}\boldsymbol{v}^\top(t_k^-)\boldsymbol{x}(t_k)-\sum_{k=0}^{p-1}\boldsymbol{v}^\top(t_k^+)\boldsymbol{x}(t_k)\\
&\qquad=\boldsymbol{v}^\top(t_p^-)\boldsymbol{x}(t_p)+\sum_{k=1}^{p-1}\Big\{\boldsymbol{v}^\top(t_k^-)-\boldsymbol{v}^\top(t_k^+)\Big\}\boldsymbol{x}(t_k)-\boldsymbol{v}^\top(t_0^+)\boldsymbol{x}(t_0). \qquad (11)
\end{aligned}$$

Substituting (11) into (10) yields

$$\begin{aligned}
J(\boldsymbol{\tau},\boldsymbol{\zeta})&=\Phi(\boldsymbol{x}(t_1),\ldots,\boldsymbol{x}(t_p),\boldsymbol{\zeta})+\sum_{k=1}^{p}\int_{t_{k-1}}^{t_k}\dot{\boldsymbol{v}}^\top(t)\boldsymbol{x}(t)dt\\
&\quad+\sum_{k=1}^{p}\int_{t_{k-1}}^{t_k}\boldsymbol{v}^\top(t)\boldsymbol{f}(\boldsymbol{x}(t),\boldsymbol{x}(t-\tau_1),\ldots,\boldsymbol{x}(t-\tau_m),\boldsymbol{\zeta})dt\\
&\quad-\boldsymbol{v}^\top(t_p^-)\boldsymbol{x}(t_p)-\sum_{k=1}^{p-1}\Big\{\boldsymbol{v}^\top(t_k^-)-\boldsymbol{v}^\top(t_k^+)\Big\}\boldsymbol{x}(t_k)+\boldsymbol{v}^\top(0^+)\boldsymbol{\phi}(0,\boldsymbol{\zeta}).
\end{aligned} \qquad (12)$$

Define the state variation with respect to τ_i as follows:

$$\boldsymbol{\Lambda}^i(t)=\frac{\partial\boldsymbol{x}(t)}{\partial\tau_i},\quad t\in[0,T].$$

If $t<\tau_l$, then $\boldsymbol{x}(t-\tau_l)=\boldsymbol{\phi}(t-\tau_l,\boldsymbol{\zeta})$, and thus

$$\frac{\partial}{\partial\tau_i}\big\{\boldsymbol{x}(t-\tau_l)\big\}=\frac{\partial}{\partial\tau_i}\big\{\boldsymbol{\phi}(t-\tau_l,\boldsymbol{\zeta})\big\}=-\delta_{li}\frac{\partial\boldsymbol{\phi}(t-\tau_l,\boldsymbol{\zeta})}{\partial t}, \qquad (13)$$

where δ_{li} denotes the Kronecker delta function. On the other hand, if $t\geq\tau_l$, then

$$\frac{\partial}{\partial\tau_i}\big\{\boldsymbol{x}(t-\tau_l)\big\}=\boldsymbol{\Lambda}^i(t-\tau_l)-\delta_{li}\dot{\boldsymbol{x}}(t-\tau_l). \qquad (14)$$

Combining (13) and (14) gives

$$\frac{\partial}{\partial \tau_i}\big\{\boldsymbol{x}(t-\tau_l)\big\} = \boldsymbol{\Lambda}^i(t-\tau_l)\chi_{[\tau_l,\infty)}(t) - \delta_{li}\boldsymbol{\psi}(t-\tau_l), \tag{15}$$

where $\chi_{[\tau_l,\infty)} : \mathbb{R} \to \mathbb{R}$ is the indicator function defined by

$$\chi_{[\tau_l,\infty)}(t) = \begin{cases} 1, & \text{if } t \geq \tau_l, \\ 0, & \text{otherwise.} \end{cases}$$

Now, in view of (15), we can differentiate (12) with respect to τ_i to obtain

$$\begin{aligned}
\frac{\partial J(\boldsymbol{\tau},\boldsymbol{\zeta})}{\partial \tau_i} &= \sum_{k=1}^{p} \frac{\partial \Phi(\boldsymbol{x}(t_1),\ldots,\boldsymbol{x}(t_p),\boldsymbol{\zeta})}{\partial \boldsymbol{x}(t_k)}\boldsymbol{\Lambda}^i(t_k) + \sum_{k=1}^{p}\int_{t_{k-1}}^{t_k} \boldsymbol{v}^\top(t)\frac{\partial \bar{\boldsymbol{f}}(t)}{\partial \boldsymbol{x}}\boldsymbol{\Lambda}^i(t)dt \\
&\quad + \sum_{k=1}^{p}\sum_{l=1}^{m}\int_{t_{k-1}}^{t_k} \boldsymbol{v}^\top(t)\frac{\partial \bar{\boldsymbol{f}}(t)}{\partial \tilde{\boldsymbol{x}}^l}\boldsymbol{\Lambda}^i(t-\tau_l)\chi_{[\tau_l,\infty)}(t)dt \\
&\quad - \sum_{k=1}^{p}\int_{t_{k-1}}^{t_k} \boldsymbol{v}^\top(t)\frac{\partial \bar{\boldsymbol{f}}(t)}{\partial \tilde{\boldsymbol{x}}^i}\boldsymbol{\psi}(t-\tau_i)dt - \boldsymbol{v}^\top(t_p^-)\boldsymbol{\Lambda}^i(t_p) \\
&\quad - \sum_{k=1}^{p-1}\Big\{\boldsymbol{v}^\top(t_k^-) - \boldsymbol{v}^\top(t_k^+)\Big\}\boldsymbol{\Lambda}^i(t_k) + \sum_{k=1}^{p}\int_{t_{k-1}}^{t_k} \dot{\boldsymbol{v}}^\top(t)\boldsymbol{\Lambda}^i(t)dt.
\end{aligned}$$

Thus,

$$\begin{aligned}
\frac{\partial J(\boldsymbol{\tau},\boldsymbol{\zeta})}{\partial \tau_i} &= \sum_{k=1}^{p-1}\Big\{\frac{\partial \Phi(\boldsymbol{x}(t_1),\ldots,\boldsymbol{x}(t_p),\boldsymbol{\zeta})}{\partial \boldsymbol{x}(t_k)} - \boldsymbol{v}^\top(t_k^-) + \boldsymbol{v}^\top(t_k^+)\Big\}\boldsymbol{\Lambda}^i(t_k) \\
&\quad - \boldsymbol{v}^\top(t_p^-)\boldsymbol{\Lambda}^i(t_p) + \frac{\partial \Phi(\boldsymbol{x}(t_1),\ldots,\boldsymbol{x}(t_p),\boldsymbol{\zeta})}{\partial \boldsymbol{x}(t_p)}\boldsymbol{\Lambda}^i(t_p) \\
&\quad + \int_0^{t_p}\Big\{\dot{\boldsymbol{v}}^\top(t) + \boldsymbol{v}^\top(t)\frac{\partial \bar{\boldsymbol{f}}(t)}{\partial \boldsymbol{x}}\Big\}\boldsymbol{\Lambda}^i(t)dt \\
&\quad + \sum_{l=1}^{m}\int_0^{t_p} \boldsymbol{v}^\top(t)\frac{\partial \bar{\boldsymbol{f}}(t)}{\partial \tilde{\boldsymbol{x}}^l}\boldsymbol{\Lambda}^i(t-\tau_l)\chi_{[\tau_l,\infty)}(t)dt \\
&\quad - \int_0^{t_p} \boldsymbol{v}^\top(t)\frac{\partial \bar{\boldsymbol{f}}(t)}{\partial \tilde{\boldsymbol{x}}^i}\boldsymbol{\psi}(t-\tau_i)dt.
\end{aligned} \tag{16}$$

Perform a change of variable in the second last integral term in (16):

$$\begin{aligned}
&\int_0^{t_p} \boldsymbol{v}^\top(t)\frac{\partial \bar{\boldsymbol{f}}(t)}{\partial \tilde{\boldsymbol{x}}^l}\boldsymbol{\Lambda}^i(t-\tau_l)\chi_{[\tau_l,\infty)}(t)dt \\
&\qquad = \int_{-\tau_l}^{t_p-\tau_l} \boldsymbol{v}^\top(t+\tau_l)\frac{\partial \bar{\boldsymbol{f}}(t+\tau_l)}{\partial \tilde{\boldsymbol{x}}^l}\boldsymbol{\Lambda}^i(t)\chi_{[0,\infty)}(t)dt \\
&\qquad = \int_{0}^{t_p-\tau_l} \boldsymbol{v}^\top(t+\tau_l)\frac{\partial \bar{\boldsymbol{f}}(t+\tau_l)}{\partial \tilde{\boldsymbol{x}}^l}\boldsymbol{\Lambda}^i(t)dt. \qquad (17)
\end{aligned}$$

Substituting (17) into (16) gives,

$$\begin{aligned}
\frac{\partial J(\boldsymbol{\tau},\boldsymbol{\zeta})}{\partial \tau_i} = &\sum_{k=1}^{p-1}\left\{\frac{\partial \Phi(\boldsymbol{x}(t_1),\ldots,\boldsymbol{x}(t_p),\boldsymbol{\zeta})}{\partial \boldsymbol{x}(t_k)} - \boldsymbol{v}^\top(t_k^-) + \boldsymbol{v}^\top(t_k^+)\right\}\boldsymbol{\Lambda}^i(t_k) \\
&+ \left\{\frac{\partial \Phi(\boldsymbol{x}(t_1),\ldots,\boldsymbol{x}(t_p),\boldsymbol{\zeta})}{\partial \boldsymbol{x}(t_p)} - \boldsymbol{v}^\top(t_p^-)\right\}\boldsymbol{\Lambda}^i(t_p) + \int_0^{t_p} \dot{\boldsymbol{v}}^\top(t)\boldsymbol{\Lambda}^i(t)dt \\
&+ \int_0^{t_p} \boldsymbol{v}^\top(t)\frac{\partial \bar{\boldsymbol{f}}(t)}{\partial \boldsymbol{x}}\boldsymbol{\Lambda}^i(t)dt + \sum_{l=1}^{m}\int_0^{t_p-\tau_l} \boldsymbol{v}^\top(t+\tau_l)\frac{\partial \bar{\boldsymbol{f}}(t+\tau_l)}{\partial \tilde{\boldsymbol{x}}^l}\boldsymbol{\Lambda}^i(t)dt \\
&- \int_0^{t_p} \boldsymbol{v}^\top(t)\frac{\partial \bar{\boldsymbol{f}}(t)}{\partial \tilde{\boldsymbol{x}}^i}\boldsymbol{\psi}(t-\tau_i)dt.
\end{aligned} \qquad (18)$$

Recall that $\boldsymbol{v}$ is arbitrary. Choosing $\boldsymbol{v} = \boldsymbol{\lambda}(\cdot|\boldsymbol{\tau},\boldsymbol{\zeta})$ and substituting (6)-(8) into (18) completes the proof. □

3.2. Gradient with respect to system parameters

We now turn our attention to the gradient of J with respect to ζ_j, $j = 1,\ldots,r$. As before, let $\boldsymbol{\lambda}(\cdot|\boldsymbol{\tau},\boldsymbol{\zeta})$ be the solution of the impulsive dynamic system (6)-(8). Furthermore, for each $j = 1,\ldots,r$, define

$$\frac{\partial \bar{\boldsymbol{f}}(t|\boldsymbol{\tau},\boldsymbol{\zeta})}{\partial \zeta_j} = \frac{\partial \boldsymbol{f}(\boldsymbol{x}(t|\boldsymbol{\tau},\boldsymbol{\zeta}),\boldsymbol{x}(t-\tau_1|\boldsymbol{\tau},\boldsymbol{\zeta}),\ldots,\boldsymbol{x}(t-\tau_m|\boldsymbol{\tau},\boldsymbol{\zeta}),\boldsymbol{\zeta})}{\partial \zeta_j}.$$

Then we have the following result.

Theorem 2. *For each* $(\boldsymbol{\tau}, \boldsymbol{\zeta}) \in \mathcal{T} \times \mathcal{Z}$,

$$\frac{\partial J(\boldsymbol{\tau}, \boldsymbol{\zeta})}{\partial \zeta_j} = \frac{\partial \Phi(\boldsymbol{x}(t_1|\boldsymbol{\tau}, \boldsymbol{\zeta}), \ldots, \boldsymbol{x}(t_p|\boldsymbol{\tau}, \boldsymbol{\zeta}), \boldsymbol{\zeta})}{\partial \zeta_j} + \int_0^{t_p} \boldsymbol{\lambda}^\top(t|\boldsymbol{\tau}, \boldsymbol{\zeta}) \frac{\partial \bar{\boldsymbol{f}}(t|\boldsymbol{\tau}, \boldsymbol{\zeta})}{\partial \zeta_j} dt$$
$$+ \boldsymbol{\lambda}^\top(0^+) \frac{\partial \boldsymbol{\phi}(0, \boldsymbol{\zeta})}{\partial \zeta_j} + \sum_{l=1}^{m} \int_{-\tau_l}^{0} \boldsymbol{\lambda}^\top(t + \tau_l|\boldsymbol{\tau}, \boldsymbol{\zeta}) \frac{\partial \bar{\boldsymbol{f}}(t + \tau_l|\boldsymbol{\tau}, \boldsymbol{\zeta})}{\partial \tilde{\boldsymbol{x}}^l} \frac{\partial \boldsymbol{\phi}(t, \boldsymbol{\zeta})}{\partial \zeta_j} dt. \tag{19}$$

Proof. Let $\boldsymbol{v}(\cdot)$ be as defined in the proof of Theorem 1. Recall from equation (12) that

$$J(\boldsymbol{\tau}, \boldsymbol{\zeta}) = \Phi(\boldsymbol{x}(t_1), \ldots, \boldsymbol{x}(t_p), \boldsymbol{\zeta}) + \sum_{k=1}^{p} \int_{t_{k-1}}^{t_k} \boldsymbol{v}^\top(t) \boldsymbol{f}(\boldsymbol{x}(t), \boldsymbol{x}(t - \tau_1), \ldots, \boldsymbol{x}(t - \tau_m), \boldsymbol{\zeta}) dt$$
$$- \boldsymbol{v}^\top(t_p^-) \boldsymbol{x}(t_p) - \sum_{k=1}^{p-1} \left\{ \boldsymbol{v}^\top(t_k^-) - \boldsymbol{v}^\top(t_k^+) \right\} \boldsymbol{x}(t_k) + \boldsymbol{v}^\top(0^+) \boldsymbol{\phi}(0, \boldsymbol{\zeta})$$
$$+ \sum_{k=1}^{p} \int_{t_{k-1}}^{t_k} \dot{\boldsymbol{v}}^\top(t) \boldsymbol{x}(t) dt, \tag{20}$$

where, as in the proof of Theorem 1, we omit the $\boldsymbol{\tau}$ and $\boldsymbol{\zeta}$ arguments for clarity.

Differentiating (20) with respect to ζ_j gives

$$\frac{\partial J(\boldsymbol{\tau}, \boldsymbol{\zeta})}{\partial \zeta_j} = \frac{\partial \Phi(\boldsymbol{x}(t_1), \ldots, \boldsymbol{x}(t_p), \boldsymbol{\zeta})}{\partial \zeta_j} + \sum_{k=1}^{p} \frac{\partial \Phi(\boldsymbol{x}(t_1), \ldots, \boldsymbol{x}(t_p), \boldsymbol{\zeta})}{\partial \boldsymbol{x}(t_k)} \frac{\partial \boldsymbol{x}(t_k)}{\partial \zeta_j}$$
$$+ \sum_{k=1}^{p} \int_{t_{k-1}}^{t_k} \boldsymbol{v}^\top(t) \frac{\partial \bar{\boldsymbol{f}}(t)}{\partial \boldsymbol{x}} \frac{\partial \boldsymbol{x}(t)}{\partial \zeta_j} dt + \sum_{k=1}^{p} \int_{t_{k-1}}^{t_k} \boldsymbol{v}^\top(t) \frac{\partial \bar{\boldsymbol{f}}(t)}{\partial \zeta_j} dt$$
$$+ \sum_{k=1}^{p} \sum_{l=1}^{m} \int_{t_{k-1}}^{t_k} \boldsymbol{v}^\top(t) \frac{\partial \bar{\boldsymbol{f}}(t)}{\partial \tilde{\boldsymbol{x}}^l} \frac{\partial \boldsymbol{x}(t - \tau_l)}{\partial \zeta_j} dt - \boldsymbol{v}^\top(t_p^-) \frac{\partial \boldsymbol{x}(t_p)}{\partial \zeta_j}$$
$$- \sum_{k=1}^{p-1} \left\{ \boldsymbol{v}^\top(t_k^-) - \boldsymbol{v}^\top(t_k^+) \right\} \frac{\partial \boldsymbol{x}(t_k)}{\partial \zeta_j} + \boldsymbol{v}^\top(0^+) \frac{\partial \boldsymbol{\phi}(0, \boldsymbol{\zeta})}{\partial \zeta_j}$$
$$+ \sum_{k=1}^{p} \int_{t_{k-1}}^{t_k} \dot{\boldsymbol{v}}^\top(t) \frac{\partial \boldsymbol{x}(t)}{\partial \zeta_j} dt.$$

Thus,

$$\begin{aligned}
\frac{\partial J(\boldsymbol{\tau},\boldsymbol{\zeta})}{\partial \zeta_j} &= \frac{\partial \Phi(\boldsymbol{x}(t_1),\dots,\boldsymbol{x}(t_p),\boldsymbol{\zeta})}{\partial \zeta_j} + \left\{\frac{\partial \Phi(\boldsymbol{x}(t_1),\dots,\boldsymbol{x}(t_p),\boldsymbol{\zeta})}{\partial \boldsymbol{x}(t_p)} - \boldsymbol{v}^\top(t_p^-)\right\}\frac{\partial \boldsymbol{x}(t_p)}{\partial \zeta_j} \\
&\quad + \sum_{k=1}^{p-1}\left\{\frac{\partial \Phi(\boldsymbol{x}(t_1),\dots,\boldsymbol{x}(t_p),\boldsymbol{\zeta})}{\partial \boldsymbol{x}(t_k)} - \boldsymbol{v}^\top(t_k^-) + \boldsymbol{v}^\top(t_k^+)\right\}\frac{\partial \boldsymbol{x}(t_k)}{\partial \zeta_j} \\
&\quad + \int_0^{t_p}\left\{\dot{\boldsymbol{v}}^\top(t) + \boldsymbol{v}^\top(t)\frac{\partial \bar{\boldsymbol{f}}(t)}{\partial \boldsymbol{x}}\right\}\frac{\partial \boldsymbol{x}(t)}{\partial \zeta_j}dt + \sum_{l=1}^{m}\int_0^{t_p}\boldsymbol{v}^\top(t)\frac{\partial \bar{\boldsymbol{f}}(t)}{\partial \tilde{\boldsymbol{x}}^l}\frac{\partial \boldsymbol{x}(t-\tau_l)}{\partial \zeta_j}dt \\
&\quad + \int_0^{t_p}\boldsymbol{v}^\top(t)\frac{\partial \bar{\boldsymbol{f}}(t)}{\partial \zeta_j}dt + \boldsymbol{v}^\top(0^+)\frac{\partial \boldsymbol{\phi}(0,\boldsymbol{\zeta})}{\partial \zeta_j}.
\end{aligned} \tag{21}$$

Perform a change of variable in the second last integral term in (21):

$$\int_0^{t_p}\boldsymbol{v}^\top(t)\frac{\partial \bar{\boldsymbol{f}}(t)}{\partial \tilde{\boldsymbol{x}}^l}\frac{\partial \boldsymbol{x}(t-\tau_l)}{\partial \zeta_j}dt = \int_{-\tau_l}^{t_p-\tau_l}\boldsymbol{v}^\top(t+\tau_l)\frac{\partial \bar{\boldsymbol{f}}(t+\tau_l)}{\partial \tilde{\boldsymbol{x}}^l}\frac{\partial \boldsymbol{x}(t)}{\partial \zeta_j}dt. \tag{22}$$

Recall that $\boldsymbol{x}(t) = \boldsymbol{\phi}(t,\boldsymbol{\zeta})$ for all $t \le \tau_l$. Hence, from (22),

$$\begin{aligned}
\int_0^{t_p}\boldsymbol{v}^\top(t)\frac{\partial \bar{\boldsymbol{f}}(t)}{\partial \tilde{\boldsymbol{x}}^l}\frac{\partial \boldsymbol{x}(t-\tau_l)}{\partial \zeta_j}dt &= \int_{-\tau_l}^{0}\boldsymbol{v}^\top(t+\tau_l)\frac{\partial \bar{\boldsymbol{f}}(t+\tau_l)}{\partial \tilde{\boldsymbol{x}}^l}\frac{\partial \boldsymbol{\phi}(t,\boldsymbol{\zeta})}{\partial \zeta_j}dt \\
&\quad + \int_0^{t_p-\tau_l}\boldsymbol{v}^\top(t+\tau_l)\frac{\partial \bar{\boldsymbol{f}}(t+\tau_l)}{\partial \tilde{\boldsymbol{x}}^l}\frac{\partial \boldsymbol{x}(t)}{\partial \zeta_j}dt.
\end{aligned} \tag{23}$$

Substituting equation (23) into (21) gives,

$$\begin{aligned}
\frac{\partial J(\boldsymbol{\tau},\boldsymbol{\zeta})}{\partial \zeta_j} &= \frac{\partial \Phi(\boldsymbol{x}(t_1),\dots,\boldsymbol{x}(t_p),\boldsymbol{\zeta})}{\partial \zeta_j} + \left\{\frac{\partial \Phi(\boldsymbol{x}(t_1),\dots,\boldsymbol{x}(t_p),\boldsymbol{\zeta})}{\partial \boldsymbol{x}(t_p)} - \boldsymbol{v}^\top(t_p^-)\right\}\frac{\partial \boldsymbol{x}(t_p)}{\partial \zeta_j} \\
&\quad + \sum_{k=1}^{p-1}\left\{\frac{\partial \Phi(\boldsymbol{x}(t_1),\dots,\boldsymbol{x}(t_p),\boldsymbol{\zeta})}{\partial \boldsymbol{x}(t_k)} - \boldsymbol{v}^\top(t_k^-) + \boldsymbol{v}^\top(t_k^+)\right\}\frac{\partial \boldsymbol{x}(t_k)}{\partial \zeta_j} \\
&\quad + \int_0^{t_p}\left\{\dot{\boldsymbol{v}}^\top(t) + \boldsymbol{v}^\top(t)\frac{\partial \bar{\boldsymbol{f}}(t)}{\partial \boldsymbol{x}}\right\}\frac{\partial \boldsymbol{x}(t)}{\partial \zeta_j}dt + \sum_{l=1}^{m}\int_{-\tau_l}^{0}\boldsymbol{v}^\top(t+\tau_l)\frac{\partial \bar{\boldsymbol{f}}(t+\tau_l)}{\partial \tilde{\boldsymbol{x}}^l}\frac{\partial \boldsymbol{\phi}(t,\boldsymbol{\zeta})}{\partial \zeta_j}dt \\
&\quad + \sum_{l=1}^{m}\int_0^{t_p-\tau_l}\boldsymbol{v}^\top(t+\tau_l)\frac{\partial \bar{\boldsymbol{f}}(t+\tau_l)}{\partial \tilde{\boldsymbol{x}}^l}\frac{\partial \boldsymbol{x}(t)}{\partial \zeta_j}dt + \int_0^{t_p}\boldsymbol{v}^\top(t)\frac{\partial \bar{\boldsymbol{f}}(t)}{\partial \zeta_j}dt + \boldsymbol{v}^\top(0^+)\frac{\partial \boldsymbol{\phi}(0,\boldsymbol{\zeta})}{\partial \zeta_j}.
\end{aligned}$$

Choosing $\boldsymbol{v} = \boldsymbol{\lambda}(\cdot|\boldsymbol{\tau},\boldsymbol{\zeta})$ and substituting (6)-(8) into the above equation completes the proof of equation (19). □

3.3. Solving Problem (P)

On the basis of Theorems 1 and 2, we now present the following algorithm for computing the cost function (5) and its gradient at a given admissible control pair $(\boldsymbol{\tau}, \boldsymbol{\zeta}) \in \mathcal{T} \times \mathcal{Z}$.

Step 1. Solve the state system (1)-(2) from $t = 0$ to $t = T$ to obtain $\boldsymbol{x}(\cdot|\boldsymbol{\tau}, \boldsymbol{\zeta})$.

Step 2. Using $\boldsymbol{x}(\cdot|\boldsymbol{\tau}, \boldsymbol{\zeta})$, solve the impulsive system (6)-(8) from $t = T$ to $t = 0$ to obtain $\boldsymbol{\lambda}(\cdot|\boldsymbol{\tau}, \boldsymbol{\zeta})$.

Step 3. Using $\boldsymbol{x}(t_k|\boldsymbol{\tau}, \boldsymbol{\zeta})$, $k = 1, \ldots, p$, compute $J(\boldsymbol{\tau}, \boldsymbol{\zeta})$ via equation (5).

Step 4. Using $\boldsymbol{x}(\cdot|\boldsymbol{\tau}, \boldsymbol{\zeta})$ and $\boldsymbol{\lambda}(\cdot|\boldsymbol{\tau}, \boldsymbol{\zeta})$, compute $\frac{\partial J(\boldsymbol{\tau}, \boldsymbol{\zeta})}{\partial \tau_i}$, $i = 1, \ldots, m$ and $\frac{\partial J(\boldsymbol{\tau}, \boldsymbol{\zeta})}{\partial \zeta_j}$, $j = 1, \ldots, r$ via equations (9) and (19).

This algorithm can be integrated with a standard gradient-based optimization method (e.g. sequential quadratic programming) to solve Problem (P) as a nonlinear programming problem. The state system (1)-(2) evolves forward in time (starting from an initial condition), while the auxiliary system (6)-(8) evolves backwards in time (starting from a terminal condition). Thus, since the state and auxiliary systems evolve in opposite directions, and the auxiliary system depends on the solution of the state system, these two systems cannot be solved simultaneously. Instead, the state system is solved first in Step 1, and then the solution of the state system is used to solve the auxiliary system in Step 2. In practice, numerical integration methods are used to solve the state and auxiliary systems. If, when solving the auxiliary system in Step 2, the value of the state vector is required at a point that does not coincide with one of the numerical integration knot points in Step 1, then an appropriate interpolation method must be used (e.g. Hermite or Lagrange interpolation). The integrals in the gradient formulae (9) and (19) can be evaluated using standard numerical quadrature rules.

4. Application to parameter identification problems

4.1. Problem formulation

Consider the dynamic model (1)-(2). Suppose that τ_i, $i = 1, \ldots, m$ and ζ_j, $j = 1, \ldots, r$ are unknown parameters that need to be identified. Furthermore, suppose that $\{(t_k, \hat{\boldsymbol{y}}^k)\}_{k=1}^{p}$ is a given set of experimental data, where $\hat{\boldsymbol{y}}^k \in \mathbb{R}^q$ is the system output observed at sample time $t = t_k$. Here, the output $\boldsymbol{y}(t) \in \mathbb{R}^q$ is assumed to be a given function of the state and model parameters:

$$\boldsymbol{y}(t) = \boldsymbol{g}(\boldsymbol{x}(t|\boldsymbol{\tau}, \boldsymbol{\zeta}), \boldsymbol{\zeta}), \quad t \in [0, T], \tag{24}$$

where $\boldsymbol{g} : \mathbb{R}^n \times \mathbb{R}^r \to \mathbb{R}^q$.

The aim is to choose appropriate values for the unknown parameters τ_i, $i = 1, \ldots, m$ and ζ_j, $j = 1, \ldots, r$ so that the predicted system output—obtained by solving (1)-(2) and (24)—best fits the experimental data. This leads to the following *parameter identification problem*:

$$\min_{(\boldsymbol{\tau},\boldsymbol{\zeta}) \in \mathcal{T} \times \mathcal{Z}} \sum_{k=1}^{p} \left| \boldsymbol{g}(\boldsymbol{x}(t_k|\boldsymbol{\tau}, \boldsymbol{\zeta}), \boldsymbol{\zeta}) - \hat{\boldsymbol{y}}^k \right|^2. \tag{25}$$

This problem is clearly a special case of Problem (P). Hence, it can be solved using the computational approach outlined in the previous section.

A similar (but less general) parameter identification problem was recently considered in reference [12]. In [12], the method proposed for computing the cost function's gradient involves solving $mn + nr + n$ differential equations. Using the algorithm in Section 3.3, only $2n$ differential equations need to be solved. Thus, our new method is ideal for online applications in which efficiency is paramount.

4.2. Example: Zinc sulphate purification

We now demonstrate the applicability of our approach to a realistic parameter identification problem. Specifically, we consider the industrial purification process described in [2, 8]. In this process, zinc powder is added to a zinc sulphate electrolyte to encourage deposition of harmful cobalt and cadmium ions. This is a key step in the production of zinc.

The concentrations of cobalt and cadmium ions in the electrolyte evolve according to the following differential equations:

$$V\dot{x}_1(t) = Qx_1^0 - Qx_1(t-\tau) - \alpha_1 u(t)x_1(t-\tau) + \beta_1 x_2(t-\tau), \tag{26}$$
$$V\dot{x}_2(t) = Qx_2^0 - Qx_2(t-\tau) - \alpha_2 v(t)x_2(t-\tau) + \beta_2 x_1(t-\tau), \tag{27}$$

and

$$x_1(t) = 3.3 \times 10^{-4}, \quad x_2(t) = 4.0 \times 10^{-3}, \quad t \leq 0, \tag{28}$$

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 amount of zinc powder added to the reaction tank. Furthermore, V is the volume of the reaction tank ($V = 400$); Q is the flux of solution ($Q = 200$); α_1 and α_2 are unknown model parameters; β_1 and β_2 are given model parameters ($\beta_1 = 16.67$, $\beta_2 = 710.7$); and x_1^0 and x_2^0 are, respectively, the concentrations of cobalt and cadmium ions at the inlet of the reaction tank ($x_1^0 = 6 \times 10^{-4}$, $x_2^0 = 9 \times 10^{-3}$).

Reference [8] considers system (26)-(28) with a given time-delay of $\tau = 2$. Here, we suppose that τ is an unknown model parameter that needs to be identified. We assume that the terminal time is $T = 8$. Furthermore, we set the input variables u and v as equal to the optimal control functions obtained in [8]:

$$u(t) = \sum_{l=1}^{8} \sigma_1^l \chi_{[\gamma_{l-1},\gamma_l)}(t), \quad t \in [0, 8], \tag{29}$$
$$v(t) = \sum_{l=1}^{8} \sigma_2^l \chi_{[\gamma_{l-1},\gamma_l)}(t), \quad t \in [0, 8], \tag{30}$$

where the switching times γ_l and the control values σ_1^l and σ_2^l, $l = 1, \ldots, 8$ are listed in Table 1.

Table 1: Control values and switching times for control functions (29) and (30).

l	1	2	3	4	5	6	7	8
γ_l	1	2	3	4	5	6	7	8
σ_1^l ($\times 10^5$)	1.08	1.57	1.24	1.56	1.59	1.43	1.25	1.25
σ_2^l ($\times 10^5$)	5.20	4.70	4.97	4.60	4.53	4.64	4.74	4.62

The system output is the concentration of cadmium ions:

$$y(t) = x_2(t). \tag{31}$$

Given system (26)-(28) and (31), and control input functions (29) and (30), our goal is to identify the model parameters α_1 and α_2 and the state-delay τ.

To generate the observed data for this parameter identification problem, we consider system (26)-(28) with the following data:

$$\tau = \hat{\tau} = 2, \quad \alpha_1 = \hat{\alpha}_1 = 7.828 \times 10^{-4}, \quad \alpha_2 = \hat{\alpha}_2 = 2.823 \times 10^{-4}.$$

The corresponding output trajectory $y(\cdot|\hat{\tau}, \hat{\alpha}_1, \hat{\alpha}_2) = x_2(\cdot|\hat{\tau}, \hat{\alpha}_1, \hat{\alpha}_2)$ acts as our reference trajectory. We define the sample times to be $t_k = k/2$, $k = 1, \dots, 16$. Thus, the observed output is

$$\hat{y}^k = x_2(t_k|\hat{\tau}, \hat{\alpha}_1, \hat{\alpha}_2), \quad k = 1, \dots, 16.$$

Our parameter identification problem is now defined as follows: Choose τ, α_1, and α_2 to minimize

$$J(\tau, \alpha_1, \alpha_2) = \sum_{k=1}^{16} \left| y(t_k|\tau, \alpha_1, \alpha_2) - \hat{y}^k \right|^2 = \sum_{l=1}^{16} \left| x_2(t_k|\tau, \alpha_1, \alpha_2) - x_2(t_k|\hat{\tau}, \hat{\alpha}_1, \hat{\alpha}_2) \right|^2$$

subject to the dynamic system (26)-(28).

This problem cannot be solved using the identification method in [12], which is only applicable when each nonlinear term in the system dynamics

Table 2: Numerical convergence of the cost values for the example in Section 4.2.

	Initial guess			Cost value at ith iteration			
Run	τ^0	α_1^0	α_2^0	$i=0$	$i=10$	$i=20$	$i=50$
1	0.0	0.0	0.0	9.264×10^{-5}	1.514×10^{-6}	3.690×10^{-9}	2.525×10^{-11}
2	0.5	0.5	0.5	7.360×10^{54}	1.905×10^{-5}	2.150×10^{-7}	3.202×10^{-13}
3	1.0	0.0	1.0	1.537×10^{20}	1.330×10^{-7}	9.813×10^{-10}	1.290×10^{-10}
4	1.0	1.0	1.0	3.392×10^{33}	2.126	3.900×10^{-3}	2.535×10^{-11}
5	3.0	1.0	1.0	8.085×10^{13}	4.841×10^{-6}	7.072×10^{-9}	8.882×10^{-11}

contains a single delay and no unknown parameters (the third term on the right-hand side of (26) violates this requirement). We solve the parameter identification problem using a Matlab program that integrates the SQP optimization method with the gradient computation algorithm described in Section 3.3. Computational results for different initial guesses are shown in Table 2. The convergence of the output trajectory for the initial guess $\tau = 3$, $\alpha_1 = 1$, and $\alpha_2 = 1$ (run 5) is displayed in Figure 1. This figure shows the output trajectory at two intermediate iterations of the optimization process, as well as the final (converged) trajectory. In Table 2 and Figure 1, τ^i, α_1^i, and α_2^i are the values of τ, α_1, and α_2 at the ith iteration of the SQP optimization process ($i = 0$ refers to the initial guess). We see from Table 2 and Figure 1 that the system trajectory converges quickly to the observed data, even when the initial trajectory is far from the reference trajectory.

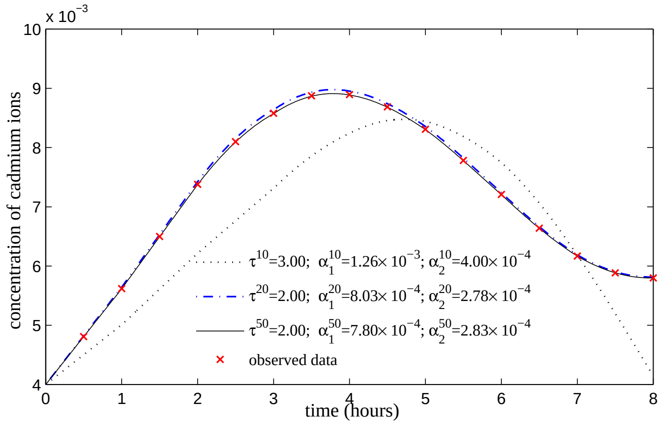

Figure 1: Numerical convergence of the output trajectory for run 5 in Section 4.2.

5. Application to delayed feedback control

5.1. Problem formulation

Consider the following continuous-time control system:

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{f}(\boldsymbol{x}(t), \boldsymbol{u}(t)), \quad t \in [0, T], \tag{32}$$

$$\boldsymbol{x}(t) = \boldsymbol{\phi}(t), \quad t \leq 0, \tag{33}$$

where $\boldsymbol{x}(t) \in \mathbb{R}^n$ is the state and $\boldsymbol{u}(t) \in \mathbb{R}^r$ is the control input. System (32)-(33) does not contain any delays. Such undelayed systems are usually much easier to control than time-delay systems. Nevertheless, it has been shown that introducing delays to an undelayed system can be beneficial, especially for chaotic systems [13, 15, 22].

Delayed feedback control is one way of deliberately introducing delays to an undelayed system. In delayed feedback control, the control function $\boldsymbol{u}(t)$ is defined as follows:

$$\boldsymbol{u}(t) = \boldsymbol{K}_0\boldsymbol{x}(t) + \boldsymbol{K}_1\boldsymbol{x}(t-\tau_1) + \cdots + \boldsymbol{K}_d\boldsymbol{x}(t-\tau_d), \tag{34}$$

where $\boldsymbol{K}_i \in \mathbb{R}^{r\times n}$, $i = 0, \ldots, d$ are feedback gain matrices and τ_i, $i = 1, \ldots, d$

are time-delays. Substituting (34) into (32)-(33) yields the following closed-loop system:

$$\dot{\boldsymbol{x}}(t) = \tilde{\boldsymbol{f}}\big(\boldsymbol{x}(t), \boldsymbol{x}(t-\tau_1), \cdots, \boldsymbol{x}(t-\tau_d), \boldsymbol{\xi}\big), \quad t \in [0, T], \tag{35}$$

$$\boldsymbol{x}(t) = \boldsymbol{\phi}(t), \quad t \leq 0, \tag{36}$$

where $\boldsymbol{\xi} \in \mathbb{R}^{rn(d+1)}$ is a vector containing the elements of the feedback gain matrices and

$$\tilde{\boldsymbol{f}}\big(\boldsymbol{x}(t), \boldsymbol{x}(t-\tau_1), \ldots, \boldsymbol{x}(t-\tau_d), \boldsymbol{\xi}\big) = \boldsymbol{f}\big(\boldsymbol{K}_0\boldsymbol{x}(t) + \boldsymbol{K}_1\boldsymbol{x}(t-\tau_1) + \ldots + \boldsymbol{K}_d\boldsymbol{x}(t-\tau_d)\big).$$

The aim here is to choose the delays and feedback gain matrices in (34) to stabilize the closed-loop system (35)-(36). Thus, we consider the following optimization problem:

$$\min_{\boldsymbol{\tau}, \boldsymbol{\xi}} |\boldsymbol{x}(T) - \boldsymbol{x}^*|^2 + |\dot{\boldsymbol{x}}(T)|^2,$$

where $\boldsymbol{x}(\cdot)$ is the solution of (35)-(36) and $\boldsymbol{x}^*$ is a desired equilibrium point. This problem can be solved effectively using the computational approach outlined in Section 3.

5.2. Example 1: Inverted pendulum

We consider the problem of controlling the position of a single-link rotational joint in robotics (a type of inverted pendulum system). The dynamics of the rotational joint are described as follows:

$$\ddot{y}(t) - \frac{g}{L}y(t) = u(t), \quad t \in [0, T], \tag{37}$$

with initial conditions

$$\dot{y}(t) = 0, \quad y(t) = 1, \quad t \leq 0, \tag{38}$$

where y denotes the angular displacement of the inverted pendulum, g is the acceleration due to gravity ($g = 9.8\text{ms}^{-2}$), L is the length of the pendulum ($L = 0.4\text{m}$), and u is the external torque force.

In the absence of velocity measurements, the inverted pendulum system is difficult to stabilize using position feedback control [22]. Thus, it is necessary to instead consider the following delayed feedback controller:

$$u(t) = ay(t - \tau_1) + by(t - \tau_2), \tag{39}$$

where τ_1 and τ_2 are position delays, and a and b are parameters. We use the same values for a and b as given in [22]:

$$a = -63.73, \quad b = 36.76. \tag{40}$$

The second-order system (37)-(38), with u defined by (39), can be easily transformed into the following system of first-order differential equations:

$$\dot{x}_1(t) = x_2, \quad t \in [0, T], \tag{41}$$

$$\dot{x}_2(t) = ax_1(t - \tau_1) + bx_1(t - \tau_2) + \frac{g}{L}x_1(t), \quad t \in [0, T], \tag{42}$$

with initial conditions

$$x_1(t) = 1, \quad x_2(t) = 0, \quad t \leq 0. \tag{43}$$

Exponential stability conditions for system (41)-(42) were established in [22]. Here, we apply the computational method described in Section 3 to determine optimal values for the position delays so that the system becomes stable at the origin. Our optimal control problem can be stated as follows: Given system (41)-(42) with initial conditions (43) and parameter values (40), choose the position delays τ_1 and τ_2 to minimize the objective function

$$J = x_1(T)^2 + x_2(T)^2, \tag{44}$$

where the terminal time T is chosen to be 20 seconds. As in Section 4.2, we solved this problem using a Matlab program that implements the computational approach described in Section 3.3. The optimal time-delays are $\tau_1 = 0.1134$ and $\tau_2 = 0.2458$. To compare, reference [22] reports optimal time-delays of $\tau_1 = 0.143$ and $\tau_2 = 0.286$. Figure 2 shows the angular displacement under our optimal feedback controller and the optimal feedback controller in [22]. Note that our controller stabilizes the system quickly with less oscillations than the controller in [22].

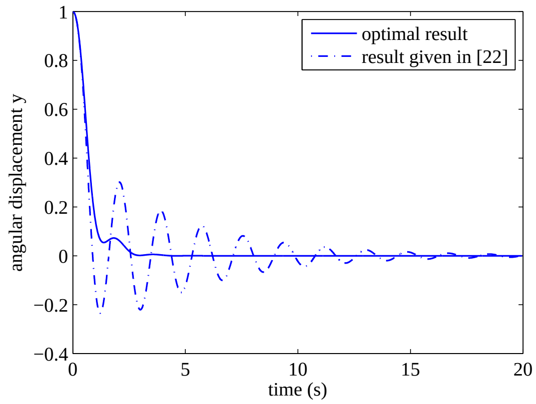

Figure 2: Optimal angular displacement for the closed-loop inverted pendulum system

234 *5.3. Example 2: Chen chaotic system*

We now consider the problem of stabilizing the so-called disturbed Chen chaotic system, which is defined as follows:

$$\dot{\boldsymbol{x}}(t) = \begin{bmatrix} -\theta_1 & \theta_1 & 0 \\ \theta_2 - \theta_1 & \theta_2 & 0 \\ 0 & 0 & -\theta_3 \end{bmatrix} \boldsymbol{x}(t) + \begin{bmatrix} 0 \\ -x_1 x_3 \\ x_1 x_2 \end{bmatrix} + \boldsymbol{\omega}(t), \quad t \in [0, T], \tag{45}$$

with initial conditions

$$\boldsymbol{x}(0) = [2, -3, 1]^\top, \quad t \leq 0, \tag{46}$$

where $\boldsymbol{\omega}(t)$ is a bounded exogenous disturbance and θ_1, θ_2, and θ_3 are model parameters. Here, we assume that the disturbance and model parameters are as given in [23]:

$$\boldsymbol{\omega}(t) = [0.2x_1(t), -0.2x_2(t), -0.2x_3(t)]^\top, \quad \theta_1 = 1, \quad \theta_2 = 2, \quad \theta_3 = 3. \tag{47}$$

Our aim is to stabilize the chaotic system (45)-(46) at the origin. Thus, the objective function is

$$J = |\boldsymbol{x}(T)|^2 + |\dot{\boldsymbol{x}}(T)|^2, \tag{48}$$

where the terminal time is $T = 0.5$. We design a delayed feedback controller in the following form:

$$\boldsymbol{u}(t) = [K_1 x_1(t-\tau), K_2 x_2(t-\tau), K_3 x_3(t-\tau)]^\top, \tag{49}$$

where K_1, K_2, K_3 are feedback gains and τ is the state-delay. Our optimal control problem can be stated as follows: Given the system (45)-(46), with disturbance and parameters values defined by (47), and the feedback control (49), choose the state-delay and the feedback gains to minimize the objective function (48).

We solved this problem using the same Matlab program that was used to solve the examples in Sections 4.2 and 5.2. The optimal delayed feedback control is

$$\boldsymbol{u}(t) = [-48.26 x_1(t-0.0071), -47.81 x_2(t-0.0071), -47.86 x_3(t-0.0071)]^\top. \tag{50}$$

Using the MISER optimal control software [24], we also computed the optimal undelayed feedback control:

$$\boldsymbol{u}(t) = [-45.47 x_1(t), -61.84 x_2(t), -20.64 x_3(t)]^\top. \tag{51}$$

The optimal state variables under controls (50) and (51) are shown in Figure 3. Note that for this system, delayed feedback control stabilizes the system quicker than the traditional feedback control.

6. Conclusion

In this paper, we have considered a novel optimal control problem in which the delays in a nonlinear time-delay system are control variables to be determined optimally. Such problems, which are called optimal state-delay control problems, arise in parameter identification and delayed feedback control. Our main contribution is a new computational method for determining the gradient of the cost function in an optimal state-delay control problem. This method requires less numerical integration than the existing method in

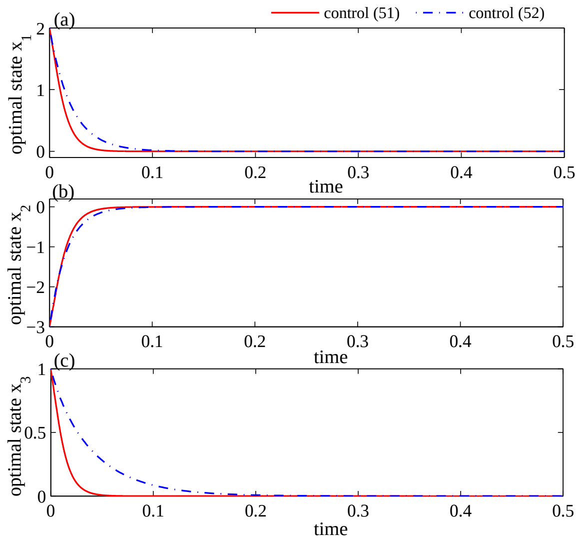

Figure 3: Optimal states of the Chen chaotic system in Section 5.3

[12], and is therefore much faster. Furthermore, unlike the method in [12], our new method is applicable to systems with nonlinear terms containing more than one state-delay. We have restricted our attention in this paper to systems with time-invariant (constant) time-delays. Our future work will involve combining the techniques in this paper with the control parameterization method [25, 26] to solve optimal state-delay control problems with time-varying delays. Such problems arise in the control of crushing processes [19] and mixing tanks with recycle loops [27].

Acknowledgment

This work was supported by the Australian Research Council (DP110100083), the China National Science Fund for Distinguished Young Scholars (61025015), and the National Natural Science Foundation of China (61174133).

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] 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(4): 705-718.

[3] L. D. Vidal, C. Jauberthie, G. J. Blanchard, Identifiability of a nonlinear delayed-differential aerospace model, IEEE Transactions on Automatic Control, 2006, 51(1): 154-158.

[4] R. F. Stengel, R. Ghigliazza, N. Kulkarni, O. Laplace, Optimal control of innate immune response, Optimal Control Applications and Methods, 2002, 23(2): 91-104.

[5] L. Göllmann, D. Kern, H. Maurer, Optimal control problems with delays in state and control variables subject to mixed control-state constraints, Optimal Control Applications and Methods, 2009, 30(4): 341-365.

[6] H. T. Banks, Necessary conditions for control problems with variable time lags, SIAM Journal on Control, 1968, 8(1): 9-47.

[7] 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.

[8] 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, in press.

[9] F. Viola, W. F. Walker, A spline-based algorithm for continuous time-delay estimation using sampled data, IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control, 2005, 52(1): 80-93.

[10] N. Salkanović, B. Lačević, B. Peruničić, Ž. Jurić, Parametric identification of plants with multiple delays and internal feedbacks using genetic algorithm, Proceddings of the 3rd International Symposium on Communications, Control and Signal Processing, 12-14, March, 2008: 425-429.

[11] Y. Orlov, L. Belkoura, J. P. Richard, M. Dambrine, Adaptive identification of linear time-delay systems, International Journal of Robust and Nonlinear Control, 2003, 13(9): 1099-1239.

[12] R. Loxton, K. L. Teo, V. Rehbock, An optimization approach to state-delay identification, IEEE Transactions on Automatic Control, 2010, 55(9): 2113-2119.

[13] A. Ahlborn, U. Parlitz, Stabilizing unstable steady states using multiple delay feedback control, Physical Review Letters, 2004,93: ID 264101.

[14] C. A. S. Batista, S. R. Lopes, R. L. Viana, A. M. Batista, Delayed feedback control of bursting synchronization in a scale-free neuronal network, Neural Networks, 2010, 23(1): 114-124.

[15] J. Lavaei, S. Sojoudi, R. M. Murray, Delay-based controller design for continuous-time and hybrid applications, Technical Report, California Institute of Technology, 2010 (available online at http://caltechcdstr.library.caltech.edu/173).

[16] N. U. Ahmed, Dynamic Systems and Control with Applications, World Scientific, Singapore, 2006.

[17] R. C. Loxton, K. L. Teo, V. Rehbock, Optimal control problems with multiple characteristic time points in the objective and constraints, Automatica, 2008, 44(11): 2923-2929.

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

[19] J. P. Richard, Time-delay systems: An overview of some recent advances and open problems, Automatica, 2003, 39(10): 1667-1694.

[20] D. G. Luenberger, Y. Ye, Linear and Nonlinear Programming, third ed., Springer, New York, 2008.

[21] J. Nocedal, S. J. Wright, Numerical Optimization, 2nd edition, Springer, 2006.

[22] D. Zhao, J. Wang, Exponential stability and spectral analysis of the inverted pendulum system under two delayed position feedbacks, Journal of Dynamical and Control Systems, 2012, 18(2): 269-295.

[23] H. Xu, Y. Chen, K. L. Teo, R. Loxton, An impulsive stabilizing control of a new chaotic system, Dynamic Systems and Applications, 2009, 18(2): 241-250.

[24] L. S. Jennings, M. E. Fisher, K. L. Teo, C.J. Goh, MISER3: Solving optimal control problems—an update, Advances in Engineering Software and Workstations, 1991, 13(4):190-196.

[25] Q. Lin, R. Loxton, K. L. Teo, Y. H. Wu, A new computational method for a class of free terminal time optimal control problems, Pacific Journal of Optimization, 2011, 7(1): 63-81.

[26] Q. Lin, R. Loxton, K. L. Teo, Y. H. Wu, Optimal control computation for nonlinear systems with state-dependent stopping criteria, Automatica, to appear.

[27] J. Y. Dieulot, J. P. Richard, Tracking control of a nonlinear system with input-dependent delay, Proceedings of the 40th IEEE Conference on Decision and Control, 2001(4): 4027-4031.