An Optimization Approach to State-Delay Identification

Ryan Loxton, Kok Lay Teo, and Volker Rehbock

Abstract—**We consider a nonlinear delay-differential system with unknown state-delays. Our goal is to identify these state-delays using experimental data. To this end, we formulate a dynamic optimization problem in which the state-delays are decision variables and the cost function measures the discrepancy between predicted and observed system output. We then show that the gradient of this problem's cost function can be computed by solving an auxiliary delay-differential system. By exploiting this result, the state-delay identification problem can be solved efficiently using a gradient-based optimization method.**

Index Terms—**Delay systems, nonlinear control systems, optimization methods, system identification.**

I. INTRODUCTION

A mathematical model for a system is typically constructed as follows. First, the system is embedded within a family of systems having a common structure, and a general model is designed to encapsulate this structure. Second, the general model is tailored to the specific system of interest by choosing appropriate values for the model parameters. This second step is called *parameter identification*. Parameter identification is usually done by comparing the system output observed in practice with the system output predicted by the model and adjusting the parameters accordingly.

In this note, we consider a parameter identification problem for a nonlinear delay-differential system. This system behaves as follows: at each time t, the system's instantaneous rate of change depends not only on its current state, but also on its state at times $t - \tau_i$, $i = 1, \ldots, m$, where each τ_i is a so-called *state-delay*. These state-delays are model parameters that need to be identified.

Naturally, we want to choose estimates for the state-delays so that the predicted system output matches the observed system output as closely as possible. Thus, in this note we formulate an optimization problem in which the state-delays are decision variables and the cost function penalizes the squared difference between predicted and observed system output. Unlike conventional optimization problems involving delay systems, the delays here are variable and influence the cost function *implicitly* through the governing delay-differential system. Thus, deriving the gradient of the cost function is a very difficult task. We will show that the cost function's gradient can be computed by solving an auxiliary delay-differential system. This is a fundamental result that enables one to solve the identification problem—and thereby obtain accurate estimates for the state-delays—using standard optimization techniques [1], [2].

We emphasize that this approach to delay identification is applicable to *nonlinear* delay-differential systems. In contrast, most other delay identification techniques are only applicable to linear systems—see, for example, [3]–[5] and the references cited therein. Yet many delay systems that arise in applications, such as predator-prey systems [6], aerospace systems [7], and continuously-stirred tank reactors [8], are actually nonlinear. Our new identification method is specifically designed to handle such systems. We demonstrate this in Section V of this note, where we apply our new method to the continuously-stirred tank reactor described in [8]. The results indicate that our method is very fast, making it ideal for online applications.

II. PROBLEM FORMULATION

Consider the following dynamic model:

$$\dot{\boldsymbol{x}}(t) = \sum_{i=1}^{m} \boldsymbol{f}^i\left(\boldsymbol{x}(t), \boldsymbol{x}(t - \tau_i)\right), \qquad t \in (0, T] \tag{1}$$

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

and

$$\boldsymbol{y}(t) = \boldsymbol{g}\left(\boldsymbol{x}(t)\right), \qquad t \leq T \tag{3}$$

where $T > 0$ is a given *terminal time*; $\boldsymbol{x}(t) \in \mathbb{R}^n$ is the *state* at time t; $\boldsymbol{y}(t) \in \mathbb{R}^r$ is the *output* at time t; τ_i, $i = 1, \ldots, m$, are unknown state-delays; and $\boldsymbol{f}^i : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$, $i = 1, \ldots, m$, $\boldsymbol{\phi} : \mathbb{R} \to \mathbb{R}^n$, and $\boldsymbol{g} : \mathbb{R}^n \to \mathbb{R}^r$ are given functions.

We assume that the following conditions are satisfied.

Assumption 1: The functions $\boldsymbol{f}^i$, $i = 1, \ldots, m$, and $\boldsymbol{g}$ are continuously differentiable, and the function $\boldsymbol{\phi}$ is twice continuously differentiable.

Assumption 2: There exists a real number $L_1 > 0$ such that for each $i = 1, \ldots, m$

$$|\boldsymbol{f}^i(\boldsymbol{w}, \boldsymbol{z})| \leq L_1\left(1 + |\boldsymbol{w}| + |\boldsymbol{z}|\right), \quad (\boldsymbol{w}, \boldsymbol{z}) \in \mathbb{R}^n \times \mathbb{R}^n$$

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

Suppose that the system modeled by (1)–(3) has been observed (for example, during an experiment) at times $t = t_j$, $j = 1, \ldots, p$. For each $j = 1, \ldots, p$, let $\hat{\boldsymbol{y}}^j \in \mathbb{R}^r$ denote the system's output measured at time $t = t_j$. Our goal is to use this experimental data to estimate the state-delays in (1)–(3).

We assume that the state-delays are non-negative and bounded above by a given real number $\bar{\tau} > 0$

$$0 \leq \tau_i \leq \bar{\tau}, \quad i = 1, \ldots, m. \tag{4}$$

A vector $\boldsymbol{\tau} \in \mathbb{R}^m$ that satisfies the constraints (4) is called a *candidate state-delay vector*. Let $\mathcal{T}$ denote the set consisting of all candidate state-delay vectors.

By [9, Theor. 3.3.3], the dynamic system (1), (2) has a unique solution corresponding to each candidate state-delay vector $\boldsymbol{\tau} \in \mathcal{T}$. We denote this solution by $\boldsymbol{x}(\cdot | \boldsymbol{\tau})$. Substituting $\boldsymbol{x}(\cdot | \boldsymbol{\tau})$ into (3) gives $\boldsymbol{y}(\cdot | \boldsymbol{\tau})$, the *predicted system output* corresponding to $\boldsymbol{\tau} \in \mathcal{T}$. That is

$$\boldsymbol{y}(t | \boldsymbol{\tau}) = \boldsymbol{g}\left(\boldsymbol{x}(t | \boldsymbol{\tau})\right), \quad t \leq T.$$

We consider the problem of choosing estimates for the state-delays so that the predicted output best fits the experimental data.

Problem P: Find a candidate state-delay vector $\tau \in \mathcal{T}$ that minimizes the following cost function:

$$J(\tau) \triangleq \sum_{j=1}^{p} |\boldsymbol{y}(t_j|\tau) - \hat{\boldsymbol{y}}^j|^2.$$

Problem P is a dynamic optimization problem with two unique characteristics. First, Problem P's cost function is not in the usual Mayer form (as a function of the final state) and instead depends on the state at multiple discrete time points. Although specialized techniques for handling this type of cost function are available (see [10]–[12]), none of them are applicable to Problem P.

Second, the decision variables in Problem P are actually the state-delays themselves. Current methods for optimizing delay systems (see, for example, [10], [13], and [14]) are only applicable when the delays are fixed and known. These methods are designed for systems of the following type:

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{f}(\boldsymbol{x}(t), \boldsymbol{x}(t - h), \zeta_1, \ldots, \zeta_q), \quad t \in (0, T]$$

where h is given and ζ_i, $i = 1, \ldots, q$, are decision variables. We are not aware of any methods for solving optimization problems like Problem P in which the delays need to be chosen optimally. The purpose of this note is to develop such a method.

We assume throughout this note that the output function $\boldsymbol{g}$ and the experimental data $\{(t_j, \hat{\boldsymbol{y}}^j)\}_{j=1}^{p}$ provide enough system information to ensure that Problem P is a sensible formulation of the state-delay identification problem (Problem P is meaningless if, for example, $p = 1$ or $\boldsymbol{g} = \boldsymbol{0}$). A thorough discussion of the precise conditions that $\boldsymbol{g}$ must satisfy is beyond the scope of this note. We instead direct the reader to [15], [16], where issues concerning delay identifiability are discussed, albeit only for linear systems.

In the dynamic model (1), (2), only the delays τ_i, $i = 1, \ldots, m$, are uncertain; all other system parameters are assumed to be known. A more realistic model should also cater for uncertain parameters in the dynamics and/or initial conditions. An example of such a model is

$$\dot{\boldsymbol{x}}(t) = \sum_{i=1}^{m} \boldsymbol{f}^i(\boldsymbol{x}(t), \boldsymbol{x}(t - \tau_i), \zeta_1, \ldots, \zeta_q), \quad t \in (0, T]$$

$$\boldsymbol{x}(t) = \boldsymbol{\phi}(t, \sigma_1, \ldots, \sigma_d), \quad\quad\quad t \leq 0$$

where ζ_i, $i = 1, \ldots, q$, σ_i, $i = 1, \ldots, d$, and τ_i, $i = 1, \ldots, m$, need to be identified. Problem P for this model involves choosing ζ_i, σ_i, and τ_i to minimize the cost function J. Since an effective method for identifying the system parameters ζ_i and σ_i has already been discussed in [10], we focus solely on identifying the state-delays τ_i in this note. In other words, we assume that any system parameters in the model are fixed and known.

III. PRELIMINARIES

In this section, we assume that $k \in \{1, \ldots, m\}$ and $\tau \in \mathcal{T}$ are arbitrary but fixed. For simplicity, we write $\boldsymbol{x}$ instead of $\boldsymbol{x}(\cdot|\tau)$.

Define

$$\mathcal{S} \triangleq \{\epsilon \in \mathbb{R} : \tau + \epsilon e^k \in \mathcal{T}\}$$

where e^k is the kth unit basis vector in $\mathbb{R}^m$. Clearly

$$\mathcal{S} = [-\tau_k, \bar{\tau} - \tau_k].$$

Thus, $\mathcal{S}$ is a closed interval of positive measure, and $0 \in \mathcal{S}$.

For each $\epsilon \in \mathcal{S}$, define three corresponding functions as follows:

$$\boldsymbol{x}^\epsilon(t) \triangleq \boldsymbol{x}(t|\tau + \epsilon e^k), \quad t \leq T,$$

$$\boldsymbol{\varphi}^\epsilon(t) \triangleq \boldsymbol{x}^\epsilon(t) - \boldsymbol{x}(t), \quad t \leq T$$

and

$$\boldsymbol{\chi}^\epsilon(t) \triangleq \begin{cases} \dot{\boldsymbol{\phi}}(t), & \text{if } t \leq 0 \\ \sum_{i=1}^{m} \boldsymbol{f}^i(\boldsymbol{x}^\epsilon(t), \boldsymbol{x}^\epsilon(t - \tau_i - \epsilon\delta_{k,i})), & \text{if } t \in (0, T] \end{cases}$$

where $\delta_{k,i}$ denotes the Kronecker delta. Furthermore, for each $\epsilon \in \mathcal{S}$ and $i = 1, \ldots, m$, define

$$\boldsymbol{\theta}^{\epsilon,i}(t) \triangleq \boldsymbol{x}^\epsilon(t - \tau_i - \epsilon\delta_{k,i}) - \boldsymbol{x}(t - \tau_i), \quad t \leq T.$$

We immediately see that

$$\boldsymbol{\theta}^{\epsilon,i}(t) = \boldsymbol{\varphi}^\epsilon(t - \tau_i), \quad t \leq T, \quad i \neq k \quad\quad (5)$$

$$\boldsymbol{\varphi}^\epsilon(t) = \boldsymbol{0}, \quad\quad\quad t \leq 0 \quad\quad (6)$$

and

$$\boldsymbol{\chi}^\epsilon(t) - \boldsymbol{\chi}^0(t) = \boldsymbol{0}, \quad t \leq 0. \quad\quad (7)$$

Furthermore, for almost all $t \in (-\infty, T]$

$$\dot{\boldsymbol{x}}^\epsilon(t) = \boldsymbol{\chi}^\epsilon(t). \quad\quad (8)$$

The proof of the following result is similar to the proof of Lemma 6.4.2 in [14].

Lemma 1: There exists a positive real number $L_2 > 0$ such that for all $\epsilon \in \mathcal{S}$

$$|\boldsymbol{x}^\epsilon(t)|, |\boldsymbol{\chi}^\epsilon(t)| \leq L_2, \quad t \in [-\bar{\tau}, T].$$

Define

$$\Psi \triangleq \{\boldsymbol{w} \in \mathbb{R}^n : |\boldsymbol{w}| \leq L_2\}.$$

Then it follows immediately from Lemma 1 that $\boldsymbol{x}^\epsilon(t) \in \Psi$ for all $t \in [-\bar{\tau}, T]$ and $\epsilon \in \mathcal{S}$.

Our next result is proved in Appendix A.

Lemma 2: There exists a positive real number $L_3 > 0$ such that for all $\epsilon \in \mathcal{S}$

$$|\boldsymbol{\varphi}^\epsilon(t)|, |\boldsymbol{\theta}^{\epsilon,k}(t)|, |\boldsymbol{\chi}^\epsilon(t) - \boldsymbol{\chi}^0(t)| \leq L_3|\epsilon|, \quad t \in [0, T].$$

Equations (6) and (7) show that $\boldsymbol{\varphi}^\epsilon$ and $\boldsymbol{\chi}^\epsilon - \boldsymbol{\chi}^0$ also satisfy the inequality in Lemma 2 for $t \leq 0$. Therefore

$$|\boldsymbol{\varphi}^\epsilon(t)|, |\boldsymbol{\chi}^\epsilon(t) - \boldsymbol{\chi}^0(t)| \leq L_3|\epsilon|, \quad t \leq T. \quad\quad (9)$$

By Lemma 2, (5), and inequality (9)

$$|\boldsymbol{\theta}^{\epsilon,i}(t)| \leq L_3|\epsilon|, \quad t \in [0, T], \quad i = 1, \ldots, m. \quad\quad (10)$$

Our final preliminary result is stated below and proved in Appendix B.

Lemma 3: For almost all $t \in [0, T]$

$$\lim_{\epsilon \to 0} \frac{\boldsymbol{\theta}^{\epsilon,k}(t) - \boldsymbol{\varphi}^\epsilon(t - \tau_k)}{\epsilon} = -\boldsymbol{\chi}^0(t - \tau_k).$$

IV. THE MAIN RESULT

For each $\tau \in \mathcal{T}$, the state trajectory $\boldsymbol{x}(\cdot|\tau)$ is a function of time. In other words, if the candidate state-delay vector is fixed, then the solution of (1), (2) is a function defined on the time horizon $(-\infty, T]$. Alternatively, we can fix $t \in (-\infty, T]$ and consider the function $\boldsymbol{x}(t|\cdot) : \mathcal{T} \to \mathbb{R}^n$ whose value at $\tau \in \mathcal{T}$ is $\boldsymbol{x}(t|\tau)$. We will show in this section

that $\boldsymbol{x}(t|\cdot)$ is differentiable on $\mathcal{T}$. This is a significant result; we will use it later to derive the gradient of the cost function J.

Let $\boldsymbol{\tau} \in \mathcal{T}$. For each $k = 1, \ldots, m$, consider the following auxiliary delay-differential system:

$$\dot{\boldsymbol{\psi}}^k(t) = \sum_{i=1}^{m} \left\{ \frac{\partial \boldsymbol{f}^i\left(\boldsymbol{x}(t), \boldsymbol{x}(t - \tau_i)\right)}{\partial \boldsymbol{x}} \boldsymbol{\psi}^k(t) \right.$$
$$+ \frac{\partial \boldsymbol{f}^i\left(\boldsymbol{x}(t), \boldsymbol{x}(t - \tau_i)\right)}{\partial \tilde{\boldsymbol{x}}} \boldsymbol{\psi}^k(t - \tau_i) \bigg\}$$
$$- \frac{\partial \boldsymbol{f}^k\left(\boldsymbol{x}(t), \boldsymbol{x}(t - \tau_k)\right)}{\partial \tilde{\boldsymbol{x}}} \boldsymbol{\chi}^0(t - \tau_k), t \in (0, T] \quad (11)$$
$$\boldsymbol{\psi}^k(t) = \mathbf{0}, \quad t \le 0 \quad (12)$$

where $\boldsymbol{x} = \boldsymbol{x}(\cdot|\boldsymbol{\tau})$, $\boldsymbol{\chi}^0$ is as defined in Section III, and $\partial/\partial\tilde{\boldsymbol{x}}$ denotes partial differentiation with respect to the delayed state. The auxiliary system (11), (12) can be solved simultaneously with the state system (1), (2). Let $\boldsymbol{\psi}^k(\cdot|\boldsymbol{\tau})$ denote the solution of (11), (12) corresponding to the candidate state-delay vector $\boldsymbol{\tau}$.

We now prove our main result.

Theorem 1: For each $t \in (0, T]$, the function $\boldsymbol{x}(t|\cdot)$ is differentiable on $\mathcal{T}$. Furthermore

$$\frac{\partial \boldsymbol{x}(t|\boldsymbol{\tau})}{\partial \tau_k} = \boldsymbol{\psi}^k(t|\boldsymbol{\tau}), \quad k = 1, \ldots, m, \quad \boldsymbol{\tau} \in \mathcal{T}.$$

Proof: Let $t \in (0, T]$, $k \in \{1, \ldots, m\}$, and $\boldsymbol{\tau} \in \mathcal{T}$ be arbitrary but fixed. To prove Theorem 1, it is sufficient to show that

$$\lim_{\epsilon \to 0} \frac{\boldsymbol{x}(t|\boldsymbol{\tau} + \epsilon e^k) - \boldsymbol{x}(t|\boldsymbol{\tau})}{\epsilon} = \boldsymbol{\psi}^k(t|\boldsymbol{\tau}). \quad (13)$$

We prove (13) in three steps.

A. Notation

Let $\mathcal{S}$, $\boldsymbol{x}$, $\boldsymbol{x}^\epsilon$, $\boldsymbol{\varphi}^\epsilon$, $\boldsymbol{\chi}^0$, and $\boldsymbol{\theta}^{\epsilon, i}$ be as defined in Section III. Furthermore, for each $\epsilon \in \mathcal{S}$, define three functions on $[0, T]$ as follows:

$$\boldsymbol{v}_1^\epsilon(s) \triangleq \sum_{i=1}^{m} \int_0^1 \left\{ \frac{\partial \boldsymbol{f}^i\left(\boldsymbol{x}(s) + \eta\boldsymbol{\varphi}^\epsilon(s), \boldsymbol{x}(s - \tau_i) + \eta\boldsymbol{\theta}^{\epsilon, i}(s)\right)}{\partial \boldsymbol{x}} \right.$$
$$\left. - \frac{\partial \boldsymbol{f}^i\left(\boldsymbol{x}(s), \boldsymbol{x}(s - \tau_i)\right)}{\partial \boldsymbol{x}} \right\} \boldsymbol{\varphi}^\epsilon(s) d\eta,$$

$$\boldsymbol{v}_2^\epsilon(s) \triangleq \sum_{i=1}^{m} \int_0^1 \left\{ \frac{\partial \boldsymbol{f}^i\left(\boldsymbol{x}(s) + \eta\boldsymbol{\varphi}^\epsilon(s), \boldsymbol{x}(s - \tau_i) + \eta\boldsymbol{\theta}^{\epsilon, i}(s)\right)}{\partial \tilde{\boldsymbol{x}}} \right.$$
$$\left. - \frac{\partial \boldsymbol{f}^i\left(\boldsymbol{x}(s), \boldsymbol{x}(s - \tau_i)\right)}{\partial \tilde{\boldsymbol{x}}} \right\} \boldsymbol{\theta}^{\epsilon, i}(s) d\eta$$

and

$$\boldsymbol{v}_3^\epsilon(s) \triangleq \frac{\partial \boldsymbol{f}^k\left(\boldsymbol{x}(s), \boldsymbol{x}(s - \tau_k)\right)}{\partial \tilde{\boldsymbol{x}}}$$
$$\times \left(\boldsymbol{\theta}^{\epsilon, k}(s) - \boldsymbol{\varphi}^\epsilon(s - \tau_k) + \epsilon\boldsymbol{\chi}^0(s - \tau_k) \right).$$

Define another function $\rho : \mathcal{S} \setminus \{0\} \to \mathbb{R}$ by

$$\rho(\epsilon) \triangleq \frac{1}{|\epsilon|} \int_0^T \left\{ |\boldsymbol{v}_1^\epsilon(s)| + |\boldsymbol{v}_2^\epsilon(s)| + |\boldsymbol{v}_3^\epsilon(s)| \right\} ds.$$

Finally, Assumption 1 implies that there exists a non-negative real number L_4 such that

$$|\partial \boldsymbol{f}^i(\boldsymbol{w}, \boldsymbol{z})/\partial \boldsymbol{x}| \le L_4, \quad (\boldsymbol{w}, \boldsymbol{z}) \in \Psi \times \Psi, \quad i = 1, \ldots, m$$
$$|\partial \boldsymbol{f}^i(\boldsymbol{w}, \boldsymbol{z})/\partial \tilde{\boldsymbol{x}}| \le L_4, \quad (\boldsymbol{w}, \boldsymbol{z}) \in \Psi \times \Psi, \quad i = 1, \ldots, m$$

where Ψ is the closed ball defined in Section III and $|\cdot|$ denotes the natural matrix norm on $\mathbb{R}^{n \times n}$.

B. Behaviour of ρ as ϵ approaches zero

We now show that $\rho \to 0$ as $\epsilon \to 0$. By Lebesgue's Dominated Convergence Theorem (see [17, ch. 5]), it is sufficient to prove the following two results:

(i) the class

$$\left\{ |\epsilon^{-1}\boldsymbol{v}_1^\epsilon| + |\epsilon^{-1}\boldsymbol{v}_2^\epsilon| + |\epsilon^{-1}\boldsymbol{v}_3^\epsilon| : \epsilon \in \mathcal{S} \setminus \{0\} \right\}$$

is equibounded on $[0, T]$;
(ii) for almost all $s \in [0, T]$

$$\lim_{\epsilon \to 0} \left\{ |\epsilon^{-1}\boldsymbol{v}_1^\epsilon(s)| + |\epsilon^{-1}\boldsymbol{v}_2^\epsilon(s)| + |\epsilon^{-1}\boldsymbol{v}_3^\epsilon(s)| \right\} = 0.$$

We prove results (i)–(ii) below.

First, since Ψ is convex, for each $\epsilon \in \mathcal{S}$ and $s \in [0, T]$,

$$\boldsymbol{x}(s) + \eta\boldsymbol{\varphi}^\epsilon(s) \in \Psi, \quad \eta \in [0, 1], \quad (14)$$
$$\boldsymbol{x}(s - \tau_i) + \eta\boldsymbol{\theta}^{\epsilon, i}(s) \in \Psi, \quad \eta \in [0, 1], \quad i = 1, \ldots, m. \quad (15)$$

Therefore, by inequalities (9) and (10)

$$|\epsilon^{-1}\boldsymbol{v}_1^\epsilon(s)| \le 2mL_3L_4, \quad s \in [0, T], \quad \epsilon \in \mathcal{S} \setminus \{0\} \quad (16)$$
$$|\epsilon^{-1}\boldsymbol{v}_2^\epsilon(s)| \le 2mL_3L_4, \quad s \in [0, T], \quad \epsilon \in \mathcal{S} \setminus \{0\}. \quad (17)$$

Similarly, by Lemma 1 and inequalities (9) and (10),

$$|\epsilon^{-1}\boldsymbol{v}_3^\epsilon(s)| \le L_4(2L_3 + L_2), \quad s \in [0, T], \quad \epsilon \in \mathcal{S} \setminus \{0\}, \quad (18)$$

where L_2 is as defined in Lemma 1. Result (i) follows immediately from inequalities (16)–(18).

Now, it is clear from inequalities (9) and (10) that the following limits exist uniformly with respect to $\eta \in [0, 1]$ and $s \in [0, T]$:

$$\lim_{\epsilon \to 0} \left\{ \boldsymbol{x}(s) + \eta\boldsymbol{\varphi}^\epsilon(s) \right\} = \boldsymbol{x}(s),$$
$$\lim_{\epsilon \to 0} \left\{ \boldsymbol{x}(s - \tau_i) + \eta\boldsymbol{\theta}^{\epsilon, i}(s) \right\} = \boldsymbol{x}(s - \tau_i).$$

Furthermore, this convergence takes place inside the closed ball Ψ [see inclusions (14) and (15)]. Also note from Assumption 1 that $\partial \boldsymbol{f}^i / \partial \boldsymbol{x}$ and $\partial \boldsymbol{f}^i / \partial \tilde{\boldsymbol{x}}$ are uniformly continuous on $\Psi \times \Psi$. Hence, the following limit exists uniformly with respect to $\eta \in [0, 1]$ and $s \in [0, T]$:

$$\lim_{\epsilon \to 0} \frac{\partial \boldsymbol{f}^i\left(\boldsymbol{x}(s) + \eta\boldsymbol{\varphi}^\epsilon(s), \boldsymbol{x}(s - \tau_i) + \eta\boldsymbol{\theta}^{\epsilon, i}(s)\right)}{\partial \boldsymbol{x}}$$
$$= \frac{\partial \boldsymbol{f}^i\left(\boldsymbol{x}(s), \boldsymbol{x}(s - \tau_i)\right)}{\partial \boldsymbol{x}}$$

and a similar limit holds for $\partial \boldsymbol{f}^i / \partial \tilde{\boldsymbol{x}}$. These two limits, together with inequalities (9) and (10), imply that $\epsilon^{-1}\boldsymbol{v}_1^\epsilon$ and $\epsilon^{-1}\boldsymbol{v}_2^\epsilon$ converge to zero uniformly on $[0, T]$ as $\epsilon \to 0$. Furthermore, Lemma 3 implies that $\epsilon^{-1}\boldsymbol{v}_3^\epsilon$ converges to zero almost everywhere on $[0, T]$ as $\epsilon \to 0$. Result (ii) then follows readily.

C. Comparing $\epsilon^{-1}\boldsymbol{\varphi}^\epsilon$ with $\boldsymbol{\psi}^k(\cdot|\boldsymbol{\tau})$

For each $\epsilon \in \mathcal{S}$, we have

$$\boldsymbol{\varphi}^\epsilon(t) = \boldsymbol{x}^\epsilon(t) - \boldsymbol{x}(t)$$
$$= \sum_{i=1}^{m} \int_0^t \left\{ \boldsymbol{f}^i\left(\boldsymbol{x}^\epsilon(s), \boldsymbol{x}^\epsilon(s - \tau_i - \epsilon\delta_{k, i})\right) \right.$$
$$\left. - \boldsymbol{f}^i\left(\boldsymbol{x}(s), \boldsymbol{x}(s - \tau_i)\right) \right\} ds.$$

Thus, by the mean value theorem

$$\boldsymbol{\varphi}^\epsilon(t) = \sum_{i=1}^{m} \int_0^t \left\{ \frac{\partial \boldsymbol{f}^i\left(\boldsymbol{x}(s), \boldsymbol{x}(s-\tau_i)\right)}{\partial \boldsymbol{x}} \boldsymbol{\varphi}^\epsilon(s) \right.$$
$$\left. + \frac{\partial \boldsymbol{f}^i\left(\boldsymbol{x}(s), \boldsymbol{x}(s-\tau_i)\right)}{\partial \tilde{\boldsymbol{x}}} \boldsymbol{\theta}^{\epsilon,i}(s) \right\} ds$$
$$+ \int_0^t \left\{ \boldsymbol{v}_1^\epsilon(s) + \boldsymbol{v}_2^\epsilon(s) \right\} ds.$$

Applying (5) gives

$$\boldsymbol{\varphi}^\epsilon(t) = \sum_{i=1}^{m} \int_0^t \left\{ \frac{\partial \boldsymbol{f}^i\left(\boldsymbol{x}(s), \boldsymbol{x}(s-\tau_i)\right)}{\partial \boldsymbol{x}} \boldsymbol{\varphi}^\epsilon(s) \right.$$
$$\left. + \frac{\partial \boldsymbol{f}^i\left(\boldsymbol{x}(s), \boldsymbol{x}(s-\tau_i)\right)}{\partial \tilde{\boldsymbol{x}}} \boldsymbol{\varphi}^\epsilon(s-\tau_i) \right\} ds$$
$$+ \int_0^t \frac{\partial \boldsymbol{f}^k\left(\boldsymbol{x}(s), \boldsymbol{x}(s-\tau_k)\right)}{\partial \tilde{\boldsymbol{x}}} \left(\boldsymbol{\theta}^{\epsilon,k}(s) - \boldsymbol{\varphi}^\epsilon(s-\tau_k) \right) ds$$
$$+ \int_0^t \left\{ \boldsymbol{v}_1^\epsilon(s) + \boldsymbol{v}_2^\epsilon(s) \right\} ds. \tag{19}$$

Now, integrating the auxiliary system (11), (12) yields

$$\boldsymbol{\psi}^k(t|\boldsymbol{\tau}) = \sum_{i=1}^{m} \int_0^t \left\{ \frac{\partial \boldsymbol{f}^i\left(\boldsymbol{x}(s), \boldsymbol{x}(s-\tau_i)\right)}{\partial \boldsymbol{x}} \boldsymbol{\psi}^k(s|\boldsymbol{\tau}) \right.$$
$$\left. + \frac{\partial \boldsymbol{f}^i\left(\boldsymbol{x}(s), \boldsymbol{x}(s-\tau_i)\right)}{\partial \tilde{\boldsymbol{x}}} \boldsymbol{\psi}^k(s-\tau_i|\boldsymbol{\tau}) \right\} ds$$
$$- \int_0^t \frac{\partial \boldsymbol{f}^k\left(\boldsymbol{x}(s), \boldsymbol{x}(s-\tau_k)\right)}{\partial \tilde{\boldsymbol{x}}} \boldsymbol{\chi}^0(s-\tau_k) ds. \tag{20}$$

Multiplying (19) by ϵ^{-1}, subtracting (20), and then taking the norm of both sides gives

$$|\epsilon^{-1}\boldsymbol{\varphi}^\epsilon(t) - \boldsymbol{\psi}^k(t|\boldsymbol{\tau})|$$
$$\leq \rho(\epsilon) + \int_0^t m L_4 |\epsilon^{-1}\boldsymbol{\varphi}^\epsilon(s) - \boldsymbol{\psi}^k(s|\boldsymbol{\tau})| ds$$
$$+ \sum_{i=1}^{m} \int_0^t L_4 |\epsilon^{-1}\boldsymbol{\varphi}^\epsilon(s-\tau_i) - \boldsymbol{\psi}^k(s-\tau_i|\boldsymbol{\tau})| ds.$$

This inequality holds for all $\epsilon \in \mathcal{S} \setminus \{0\}$. Using (6) and (12), we can simplify it as follows:

$$|\epsilon^{-1}\boldsymbol{\varphi}^\epsilon(t) - \boldsymbol{\psi}^k(t|\boldsymbol{\tau})| \leq \rho(\epsilon) + \int_0^t 2m L_4 |\epsilon^{-1}\boldsymbol{\varphi}^\epsilon(s) - \boldsymbol{\psi}^k(s|\boldsymbol{\tau})| ds.$$

Applying Gronwall's Lemma (see Chapter 2 of [9]) yields

$$|\epsilon^{-1}\boldsymbol{\varphi}^\epsilon(t) - \boldsymbol{\psi}^k(t|\boldsymbol{\tau})| \leq \rho(\epsilon) \exp(2m L_4 T).$$

Thus, by taking the limit as $\epsilon \to 0$,

$$\lim_{\epsilon \to 0} |\epsilon^{-1}\boldsymbol{\varphi}^\epsilon(t) - \boldsymbol{\psi}^k(t|\boldsymbol{\tau})| \leq \exp(2m L_4 T) \lim_{\epsilon \to 0} \rho(\epsilon) = 0.$$

(Recall that $\rho(\epsilon) \to 0$ as $\epsilon \to 0$.) This completes the proof of (13). ∎

It follows immediately from Theorem 1 that for each $j = 1, \ldots, p$

$$\frac{\partial \boldsymbol{x}(t_j|\boldsymbol{\tau})}{\partial \tau_k} = \boldsymbol{\psi}^k(t_j|\boldsymbol{\tau}), \quad k = 1, \ldots, m, \quad \boldsymbol{\tau} \in \mathcal{T}.$$

Therefore, for each $\boldsymbol{\tau} \in \mathcal{T}$

$$\frac{\partial J(\boldsymbol{\tau})}{\partial \tau_k} = 2 \sum_{j=1}^{p} \left(\boldsymbol{y}(t_j|\boldsymbol{\tau}) - \hat{\boldsymbol{y}}^j \right)^T \frac{\partial \boldsymbol{g}\left(\boldsymbol{x}(t_j|\boldsymbol{\tau})\right)}{\partial \boldsymbol{x}} \boldsymbol{\psi}^k(t_j|\boldsymbol{\tau}). \tag{21}$$

We now present the following algorithm for computing the cost $J(\boldsymbol{\tau})$ and cost gradient $\partial J(\boldsymbol{\tau})/\partial \boldsymbol{\tau}$ corresponding to a candidate state-delay vector $\boldsymbol{\tau} \in \mathcal{T}$.

(i) Obtain $\boldsymbol{x}(t_j|\boldsymbol{\tau})$ and $\boldsymbol{\psi}^k(t_j|\boldsymbol{\tau})$, $j = 1, \ldots, p$, by solving the delay-differential system consisting of (1), (2) and (11), (12).
(ii) For each $j = 1, \ldots, p$, use $\boldsymbol{x}(t_j|\boldsymbol{\tau})$ to compute $\boldsymbol{y}(t_j|\boldsymbol{\tau})$.
(iii) Use $\boldsymbol{y}(t_j|\boldsymbol{\tau})$, $j = 1, \ldots, p$, to compute $J(\boldsymbol{\tau})$.
(iv) Use $\boldsymbol{x}(t_j|\boldsymbol{\tau})$, $\boldsymbol{y}(t_j|\boldsymbol{\tau})$, and $\boldsymbol{\psi}^k(t_j|\boldsymbol{\tau})$ to compute $\partial J(\boldsymbol{\tau})/\partial \tau_k$, $k = 1, \ldots, m$, according to (21).

This algorithm can be readily incorporated into a standard gradient-based optimization method, such as a conjugate gradient method (see [1] and [2]). Thus, by invoking this algorithm, we can treat Problem P as a standard optimization problem and solve it using existing techniques. This yields state-delay estimates that minimize the discrepancy between predicted and observed system output. In the next section, we apply this approach to an example.

V. AN EXAMPLE

Consider a continuously-stirred tank reactor in which the chemical reaction $A \to B$ occurs. A dynamic model for this system is [8]

$$\dot{x}_1(t) = -2x_1(t) + \frac{1}{10}\left(1 - x_1(t)\right) \exp\left[\frac{20 x_2(t)}{20 + x_2(t)}\right]$$
$$+ x_1(t-\tau) \tag{22}$$
$$\dot{x}_2(t) = -\frac{5}{2}x_2(t) + \frac{4}{5}\left(1 - x_1(t)\right) \exp\left[\frac{20 x_2(t)}{20 + x_2(t)}\right]$$
$$+ x_2(t-\tau) \tag{23}$$

and

$$x_1(t) = 1, \quad x_2(t) = 1, \quad t \leq 0 \tag{24}$$

where $x_1(t)$ is the concentration of A at time t (dimensionless), $x_2(t)$ is the temperature of the reactor at time t (dimensionless), and τ is a state-delay that needs to be identified. We assume that the terminal time here is $T = 10$. We also assume that the reactor's temperature is the only quantity that can be measured. Thus, the output is

$$y(t) = x_2(t), \quad t \leq 10.$$

We use the output trajectory of (22)–(24) with $\tau = 2$ to generate the observed data in Problem P. More specifically, we set $t_j \triangleq j/2$ and $\hat{y}^j \triangleq x_2(t_j|2)$, $j = 1, \ldots, 20$. Thus, our state-delay identification problem is: choose τ to minimize

$$J(\tau) = \sum_{j=1}^{20} |y(t_j|\tau) - \hat{y}^j|^2 = \sum_{j=1}^{20} |x_2(t_j|\tau) - x_2(t_j|2)|^2$$

subject to the dynamics (22)–(24).

We solved this problem using a Fortran program that combines the gradient computation algorithm described in Section IV with the optimization routine NLPQLP (see [18]). This program uses LSODA (see [19]) to solve the state and auxiliary systems. Computational results for initial guesses of $\tau = 1$ and $\tau = 3$ are summarized in Table I and displayed in Figs. 1 and 2. Note that our program only needed eight

TABLE I
STATE-DELAY ESTIMATE AT THE iTH ITERATION.
(a) Initial guess $\tau = 1$. (b) Initial guess $\tau = 3$

(a)

i	τ^i	$J(\tau^i)$
1	1.0000	4.4918×10^{-1}
2	2.3580	1.9874×10^{-2}
3	2.2655	1.1468×10^{-2}
4	1.8883	2.5747×10^{-3}
5	2.0294	1.5965×10^{-4}
6	2.0037	2.5989×10^{-6}
7	1.9999	2.1848×10^{-9}
8	2.0000	3.0413×10^{-14}

(b)

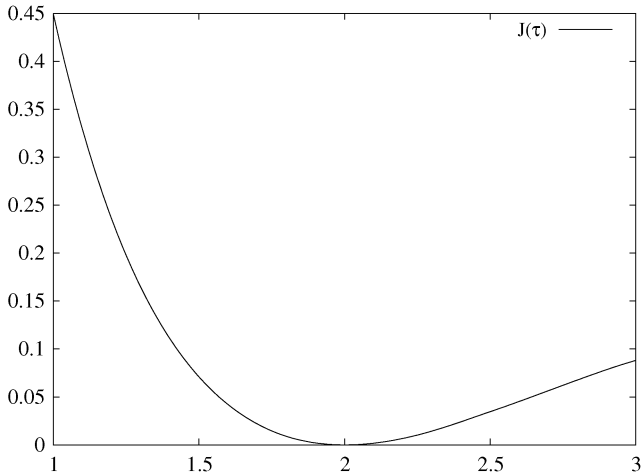| i	τ^i	$J(\tau^i)$
1	3.0000	8.8143×10^{-2}
2	2.9133	7.9892×10^{-2}
3	2.4032	2.4488×10^{-2}
4	2.1632	4.5562×10^{-3}
5	1.9110	1.6018×10^{-3}
6	2.0148	4.1125×10^{-5}
7	2.0015	4.2900×10^{-7}
8	2.0000	9.3648×10^{-11}

Fig. 1. Cost function J in the example.

iterations to converge from either initial guess to the optimal solution $\tau = 2$.

APPENDIX A
PROOF OF LEMMA 2

Let $\epsilon \in \mathcal{S}$ be arbitrary but fixed. For each $s \in [0, T]$

$$|\boldsymbol{\theta}^{\epsilon,k}(s)| = |\boldsymbol{x}^\epsilon(s - \tau_k - \epsilon) - \boldsymbol{x}(s - \tau_k)|$$
$$\leq |\boldsymbol{x}^\epsilon(s - \tau_k - \epsilon) - \boldsymbol{x}(s - \tau_k)| + |\boldsymbol{\varphi}^\epsilon(s - \tau_k)|.$$

Hence, by (8),

$$|\boldsymbol{\theta}^{\epsilon,k}(s)| \leq \int_{a(s)}^{b(s)} |\boldsymbol{\chi}^\epsilon(\eta)| d\eta + |\boldsymbol{\varphi}^\epsilon(s - \tau_k)|, \quad s \in [0, T] \quad (25)$$

where

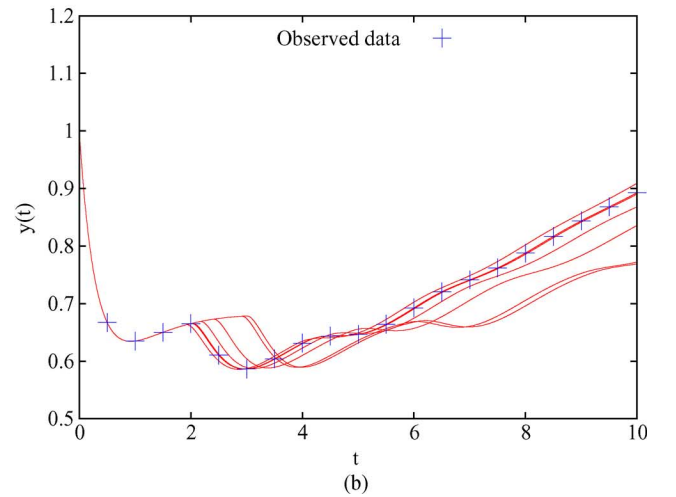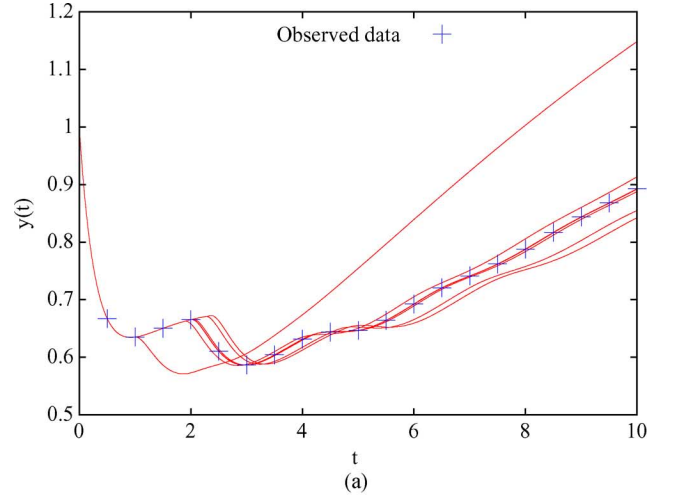$$a(s) \triangleq \min\{s - \tau_k, s - \tau_k - \epsilon\}$$

Fig. 2. Numerical convergence of the output trajectory: the red lines show $y(t) = x_2(t)$ at each iteration; the blue crosses show the observed data (which corresponds to $\tau = 2$). (a) Initial guess $\tau = 1$. (b) Initial guess $\tau = 3$.

and

$$b(s) \triangleq \max\{s - \tau_k, s - \tau_k - \epsilon\}.$$

Clearly

$$b(s) - a(s) = |\epsilon|, \quad s \in [0, T] \quad (26)$$

and

$$[a(s), b(s)] \subset [-\bar{\tau}, T], \quad s \in [0, T]. \quad (27)$$

Simplifying (25) using (26), (27), and Lemma 1 gives

$$|\boldsymbol{\theta}^{\epsilon,k}(s)| \leq L_2|\epsilon| + |\boldsymbol{\varphi}^\epsilon(s - \tau_k)|, \quad s \in [0, T]. \quad (28)$$

Now

$$|\boldsymbol{\chi}^\epsilon(s) - \boldsymbol{\chi}^0(s)| \leq \sum_{i=1}^m |\boldsymbol{f}^i(\boldsymbol{x}^\epsilon(s), \boldsymbol{x}^\epsilon(s - \tau_i - \epsilon\delta_{k,i}))$$
$$- \boldsymbol{f}^i(\boldsymbol{x}(s), \boldsymbol{x}(s - \tau_i))|, \quad s \in [0, T].$$

By Assumption 1, the functions $\boldsymbol{f}^i$, $i = 1, \ldots, m$, are Lipschitz continuous on $\Psi \times \Psi$. Hence, there exists a real number $\alpha_1 > 0$ such that

$$
\begin{aligned}
|\boldsymbol{\chi}^\epsilon(s) - \boldsymbol{\chi}^0(s)| &\leq m\alpha_1 |\boldsymbol{\varphi}^\epsilon(s)| + \sum_{i=1}^m \alpha_1 |\boldsymbol{\theta}^{\epsilon,i}(s)| \\
&= m\alpha_1 |\boldsymbol{\varphi}^\epsilon(s)| + \alpha_1 |\boldsymbol{\theta}^{\epsilon,k}(s)| \\
&\quad + \sum_{\substack{i=1 \\ i \neq k}}^m \alpha_1 |\boldsymbol{\varphi}^\epsilon(s - \tau_i)|, \quad s \in [0, T]
\end{aligned} \tag{29}
$$

where we used (5) to simplify the right-hand side. Substituting (28) into (29) gives

$$
\begin{aligned}
|\boldsymbol{\chi}^\epsilon(s) - \boldsymbol{\chi}^0(s)| &\leq \alpha_1 L_2 |\epsilon| + m\alpha_1 |\boldsymbol{\varphi}^\epsilon(s)| \\
&\quad + \sum_{i=1}^m \alpha_1 |\boldsymbol{\varphi}^\epsilon(s - \tau_i)|, \quad s \in [0, T]. \tag{30}
\end{aligned}
$$

Now, if $t \in [0, T]$, then it follows from (8) that

$$
|\boldsymbol{\varphi}^\epsilon(t)| = |\boldsymbol{x}^\epsilon(t) - \boldsymbol{x}(t)| \leq \int_0^t |\boldsymbol{\chi}^\epsilon(s) - \boldsymbol{\chi}^0(s)| ds.
$$

Applying inequality (30) yields

$$
\begin{aligned}
|\boldsymbol{\varphi}^\epsilon(t)| &\leq \alpha_1 L_2 T |\epsilon| + \int_0^t m\alpha_1 |\boldsymbol{\varphi}^\epsilon(s)| ds \\
&\quad + \sum_{i=1}^m \int_0^t \alpha_1 |\boldsymbol{\varphi}^\epsilon(s - \tau_i)| ds.
\end{aligned}
$$

Hence

$$
\begin{aligned}
|\boldsymbol{\varphi}^\epsilon(t)| &\leq \alpha_1 L_2 T |\epsilon| + \int_0^t m\alpha_1 |\boldsymbol{\varphi}^\epsilon(s)| ds \\
&\quad + \sum_{i=1}^m \int_{-\tau_i}^{t - \tau_i} \alpha_1 |\boldsymbol{\varphi}^\epsilon(s)| ds \\
&\leq \alpha_1 L_2 T |\epsilon| + \int_0^t m\alpha_1 |\boldsymbol{\varphi}^\epsilon(s)| ds \\
&\quad + \int_{-\bar{\tau}}^t m\alpha_1 |\boldsymbol{\varphi}^\epsilon(s)| ds.
\end{aligned}
$$

Recall that $\boldsymbol{\varphi}^\epsilon(s) = \boldsymbol{0}$ for all $s \leq 0$ [see (6)]. Therefore

$$
|\boldsymbol{\varphi}^\epsilon(t)| \leq \alpha_1 L_2 T |\epsilon| + \int_0^t 2m\alpha_1 |\boldsymbol{\varphi}^\epsilon(s)| ds.
$$

Finally, by Gronwall's Lemma (see [9, ch. 2])

$$
|\boldsymbol{\varphi}^\epsilon(t)| \leq \alpha_2 |\epsilon|, \quad t \in [0, T],
$$

where

$$
\alpha_2 \triangleq \alpha_1 L_2 T \exp(2m\alpha_1 T).
$$

Actually, because of (6)

$$
|\boldsymbol{\varphi}^\epsilon(t)| \leq \alpha_2 |\epsilon|, \quad t \leq T. \tag{31}
$$

Substituting (31) into (28) and (30) completes the proof.

APPENDIX B
PROOF OF LEMMA 3

We will prove Lemma 3 by showing that

$$
\lim_{\epsilon \to 0} \frac{\boldsymbol{\theta}^{\epsilon,k}(t) - \boldsymbol{\varphi}^\epsilon(t - \tau_k)}{\epsilon} = -\boldsymbol{\chi}^0(t - \tau_k), \quad t \in [0, T] \setminus \{\tau_k\}. \tag{32}
$$

First, let $t \in [0, T] \setminus \{\tau_k\}$ be arbitrary but fixed. Then for each $\epsilon \in \mathcal{S} \setminus \{0\}$

$$
\boldsymbol{\theta}^{\epsilon,k}(t) - \boldsymbol{\varphi}^\epsilon(t - \tau_k) = \boldsymbol{x}^\epsilon(t - \tau_k - \epsilon) - \boldsymbol{x}^\epsilon(t - \tau_k).
$$

Hence, by (8)

$$
\frac{\boldsymbol{\theta}^{\epsilon,k}(t) - \boldsymbol{\varphi}^\epsilon(t - \tau_k)}{\epsilon} = \frac{1}{\epsilon} \int_{t - \tau_k}^{t - \tau_k - \epsilon} \boldsymbol{\chi}^\epsilon(s) ds.
$$

We can rewrite this equation as follows:

$$
\frac{\boldsymbol{\theta}^{\epsilon,k}(t) - \boldsymbol{\varphi}^\epsilon(t - \tau_k)}{\epsilon} = -\boldsymbol{\chi}^0(t - \tau_k) + \boldsymbol{\omega}(\epsilon) \tag{33}
$$

where

$$
\boldsymbol{\omega}(\epsilon) \triangleq \frac{1}{\epsilon} \int_{t - \tau_k}^{t - \tau_k - \epsilon} \left\{ \boldsymbol{\chi}^\epsilon(s) - \boldsymbol{\chi}^0(t - \tau_k) \right\} ds.
$$

We now complete the proof by showing that $\boldsymbol{\omega}(\epsilon) \to 0$ as $\epsilon \to 0$ [(32) then follows immediately from (33)].

By the triangle inequality

$$
\begin{aligned}
|\boldsymbol{\omega}(\epsilon)| &\leq \frac{1}{|\epsilon|} \int_{a_\epsilon}^{b_\epsilon} |\boldsymbol{\chi}^\epsilon(s) - \boldsymbol{\chi}^0(s)| ds \\
&\quad + \frac{1}{|\epsilon|} \int_{a_\epsilon}^{b_\epsilon} |\boldsymbol{\chi}^0(s) - \boldsymbol{\chi}^0(t - \tau_k)| ds \tag{34}
\end{aligned}
$$

where

$$
a_\epsilon \triangleq \min\{t - \tau_k - \epsilon, t - \tau_k\}
$$

and

$$
b_\epsilon \triangleq \max\{t - \tau_k - \epsilon, t - \tau_k\}.
$$

Clearly

$$
b_\epsilon - a_\epsilon = |\epsilon|. \tag{35}
$$

Using (9) and (35) to simplify the first integral in (34) gives

$$
|\boldsymbol{\omega}(\epsilon)| \leq L_3 |\epsilon| + \frac{1}{|\epsilon|} \int_{a_\epsilon}^{b_\epsilon} |\boldsymbol{\chi}^0(s) - \boldsymbol{\chi}^0(t - \tau_k)| ds. \tag{36}
$$

This inequality holds for every $\epsilon \in \mathcal{S} \setminus \{0\}$.

Now, since $t \neq \tau_k$, either $t > \tau_k$ or $t < \tau_k$. We assume that $t > \tau_k$ (the proof when $t < \tau_k$ is similar). Accordingly, the following implication holds:

$$
\epsilon \in \mathcal{S}, |\epsilon| < t - \tau_k \implies [a_\epsilon, b_\epsilon] \subset (0, T). \tag{37}
$$

Recall the derivation of inequality (30) in Appendix A. Using similar ideas, one can show that $\boldsymbol{\chi}^0$ is Lipschitz continuous on $(0, T]$. Hence, there exists a real number $\beta_1 > 0$ such that

$$|\boldsymbol{\chi}^0(s) - \boldsymbol{\chi}^0(t - \tau_k)| \leq \beta_1 |s - t + \tau_k|, \quad s \in (0, T]. \qquad (38)$$

It follows from (37) and (38) that when the magnitude of $\epsilon \in \mathcal{S}$ is sufficiently small,

$$|\boldsymbol{\chi}^0(s) - \boldsymbol{\chi}^0(t - \tau_k)| \leq \beta_1 |s - t + \tau_k| \leq \beta_1 (b_\epsilon - a_\epsilon)$$
$$= \beta_1 |\epsilon|, \quad s \in [a_\epsilon, b_\epsilon].$$

Substituting this inequality into (36) gives

$$|\boldsymbol{\omega}(\epsilon)| \leq (L_3 + \beta_1)|\varepsilon|,$$

which holds for all $\epsilon \in \mathcal{S} \setminus \{0\}$ of sufficiently small magnitude. This shows that $\boldsymbol{\omega}(\epsilon) \to 0$ as $\epsilon \to 0$, as required.

REFERENCES

[1] J. Nocedal and S. J. Wright, *Numerical Optimization*, 2nd ed. New York: Springer, 2006.

[2] D. G. Luenberger and Y. Ye, *Linear and Nonlinear Programming*, 3rd ed. New York: Springer, 2008.

[3] Y. Orlov, L. Belkoura, J. P. Richard, and M. Dambrine, "Adaptive identification of linear time-delay systems," *Int. J. Robust Nonlin. Control*, vol. 13, no. 9, pp. 857–872, 2003.

[4] Y. Orlov, L. Belkoura, J. P. Richard, and M. Dambrine, "On identifiability of linear time-delay systems," *IEEE Trans. Autom. Control*, vol. 47, no. 8, pp. 1319–1324, Aug. 2002.

[5] J. Tuch, A. Feuer, and Z. J. Palmor, "Time delay estimation in continuous linear time-invariant systems," *IEEE Trans. Autom. Control*, vol. 39, no. 4, pp. 823–827, Apr. 1994.

[6] R. Xu and L. Chen, "Persistence and stability for a two-species ratio-dependent predator-prey system with time delay in a two-patch environment," *Comput. Math. Applic.*, vol. 40, no. 4–5, pp. 577–588, 2000.

[7] L. D. Vidal, C. Jauberthie, and G. J. Blanchard, "Identifiability of a nonlinear delayed-differential aerospace model," *IEEE Trans. Autom. Control*, vol. 51, no. 1, pp. 154–158, Jan. 2006.

[8] Y. Y. Cao and P. M. Frank, "Analysis and synthesis of nonlinear time-delay systems via fuzzy control approach," *IEEE Trans. Fuzzy Syst.*, vol. 8, no. 2, pp. 200–211, Apr. 2000.

[9] N. U. Ahmed, *Dynamic Systems and Control With Applications*. Singapore: World Scientific, 2006.

[10] L. Y. Wang, W. H. Gui, K. L. Teo, R. C. Loxton, and C. H. Yang, "Time delayed optimal control problems with multiple characteristic time points: Computation and industrial applications," *J. Indust. Manag. Optimiz.*, vol. 5, no. 4, pp. 705–718, 2009.

[11] R. Loxton, K. L. Teo, and V. Rehbock, "Optimal control problems with multiple characteristic time points in the objective and constraints," *Automatica*, vol. 44, no. 11, pp. 2923–2929, 2008.

[12] R. B. Martin, "Optimal control drug scheduling of cancer chemotherapy," *Automatica*, vol. 28, no. 6, pp. 1113–1123, 1992.

[13] L. Hasdorff, *Gradient Optimization and Nonlinear Control*. New York: Wiley, 1976.

[14] K. L. Teo, C. J. Goh, and K. H. Wong, *A Unified Computational Approach to Optimal Control Problems*. Essex, U.K.: Longman, 1991.

[15] L. Belkoura, "Identifiability of systems described by convolution equations," *Automatica*, vol. 41, no. 3, pp. 505–512, 2005.

[16] S. M. Verduyn Lunel, "Parameter identifiability of differential delay equations," *Int. J. Adapt. Control Signal Process.*, vol. 15, no. 6, pp. 655–678, 2001.

[17] R. G. Bartle, *The Elements of Integration and Lebesgue Measure*. New York: Wiley, 1995.

[18] K. Schittkowski, *NLPQLP: A Fortran Implementation of a Sequential Quadratic Programming Algorithm With Distributed and Non-Monotone Line Search*. Bayreuth, Germany: Univ. Bayreuth, 2007.

[19] A. C. Hindmarsh, "Large ordinary differential equation systems and software," *IEEE Control Syst. Mag.*, vol. 2, no. 4, pp. 24–30, Dec. 1982.

H_∞ Controller Design Using an Alternative to Youla Parameterization

Hamid Khatibi and Alireza Karimi

Abstract—All H_∞ controllers of a SISO LTI system are parameterized thanks to the relation between Bounded Real Lemma and Positive Real Lemma. This new parameterization shares the same features with Youla parameterization, namely on the convexity of H_∞ norm constraints for the closed-loop transfer functions. However, it can deal with low-order controllers and can be extended easily for the systems with polytopic uncertainty. The effectiveness of the proposed method is shown via an academic example.

Index Terms—H_∞ control, linear matrix inequalities (LMIs), robust control.

I. INTRODUCTION

Youla parameterization [1] is probably the most well-known controller parameterization, which parameterizes all stabilizing controllers of a system, over an infinite dimensional space. The main advantage of this parameterization is that all closed-loop sensitivity functions are affine w.r.t. the so-called Q parameter and hence, it can be employed for H_∞ controller design in a convex optimization problem. For this purpose, Q is defined as a linearly parameterized transfer function with fixed stable denominator and the controller parameters are computed based on the optimal Q and the plant model parameters [2]. This approach has some drawbacks.

1) The choice of the denominator of Q is crucial and may lead to conservative results.
2) The order of the controller depends on the order of the plant model and that of Q. Therefore, fixed-order controller design with an order less than that of the plant model by convex optimization is not possible.
3) The parametric uncertainty cannot be directly considered.

These problems have already been studied and some solutions been proposed that are explained hereafter. In [3], the numerator and denominator of Q are parameterized simultaneously as a linear combination of two other polynomials. The stability of Q is guaranteed by adding a strictly positive real (SPR) condition on the ratio of the denominator of Q and a "central polynomial". In this approach, fixed-order controller can be designed but parametric uncertainty is not considered.

A synthesis method is proposed in [4], which can consider norm bounded parameter uncertainty. This includes ellipsoidal parameter uncertainty and can cover the polytopic uncertainty with some conservatism. Besides, since a Q-parameterization method is involved in the synthesis approach, fixed-order controller design cannot be handled.

The problem of fixed-order H_∞ controller design by direct parameterization of controller is studied in [5] and in [6]. The infinity norm constraint on the weighted closed-loop transfer function is transformed