Electronic Research Archive

https://www.aimspress.com/journal/era

ERA, 32(10): 5889–5913.
DOI: 10.3934/era.2024272
Received: 19 August 2024
Revised: 04 October 2024
Accepted: 14 October 2024
Published: 29 October 2024

Research article

Distributionally robust parameter estimation for nonlinear fed-batch switched time-delay system with moment constraints of uncertain measured output data

Sida Lin[1], Jinlong Yuan[1,*], Zichao Liu[1], Tao Zhou[2], An Li[3], Chuanye Gu[4], Kuikui Gao[5] and Jun Xie[6]

[1] School of Science, Dalian Maritime University, Dalian 116026, China

[2] Institute for Intelligent Systems Research and Innovation (IISRI), Deakin University, Geelong, VIC, 3217, Australia

[3] School of Mathematical Sciences, Xiamen University, Xiamen 361005, China

[4] School of Management, Guangzhou University, Guangzhou 510006, China

[5] Ayata Inc., 2700 Post Oak Blvd, 21st Floor Houston, TX 77056, USA

[6] Teaching and Research Office of Mathematics, Department of Basics, PLA Dalian Naval Academy, Dalian 116018, Liaoning, China

* **Correspondence:** Email: yuanjinlong0613@163.com, yuanjinlong@dlmu.edu.cn.

Abstract: In this paper, we investigated a nonlinear continuous-time switched time-delay (NCTSTD) system for glycerol fed-batch bioconversion to 1,3-propanediol with unknown time-delay and system parameters. The measured output data was uncertain, while the first moment information about its distribution was available. Our goal was to identify these unknown quantities under the environment of uncertain measurement output data. A distributionally robust parameter estimation problem (i.e., a bi-level parameter estimation (BLPE) problem) subject to the NCTSTD system was presented, where the expectation of the discrepancy between the output of the NCTSTD system and the uncertain measured output data with respect to its probability distributions was included in the cost functional. By applying the duality theory, the BLPE problem was transformed into a single-level parameter estimation (SLPE) problem with non-smooth term approximated by a smoothing technique and its error analysis was given. Then, the gradients of the cost function of the SLPE problem were derived. A hybrid optimization algorithm was proposed for solving the SLPE problem. The paper concluded by presenting the simulation results.

Keywords: distributionally robust parameter estimation; uncertain measured output data; continuous-time switched time-delay system; error analysis; hybrid algorithm

1. Introduction

Distributionally robust optimization (DRO) is an optimization method designed to address uncertainty, particularly in situations where the underlying probability distribution is not fully known. Unlike traditional robust optimization, DRO assumes that we only have partial information about the uncertain probability distribution, rather than complete certainty. By accounting for the worst-case scenarios, DRO can reduce decision-making risks and enhance the robustness of systems across various real-world applications. It has aroused much interest among researchers to develop algorithms to solve the DRO problems in kernel density estimation [1], data-driven problem [2, 3], portfolio optimization problem [4], decision bounding problem [5], and matrix moment constraint [6]. However, no dynamical systems are involved in the aforementioned DRO problems. Recently, the DRO problems involving linear discrete-time systems are discussed in multi-access space-time block coded MIMO system [7], stochastic model predictive control problem [8], modeling of optimal control problem [9], constrained stochastic system [10], partially observable linear stochastic system [11], and output feedback [12]. However, in these literatures, the dynamical systems involved in the DRO problems are discrete-time dynamical (DTD) systems. Continuous-time dynamical (CTD) systems are not involved. There are major difficulties in solving DRO problems governed by CTD systems such as (i) a DRO problem constrained by a CTD system is fundamentally a bi-level optimization problem, which, unlike a DTD system, cannot be simplified to a single-level optimization problem with a finite number of constraints; and (ii) the presence of CTD systems in DRO problems causes the inner-level objective functional to be expressed as the expectation of a non-convex function. In this paper, we examine a DRO problem influenced by a nonlinear continuous-time switched time-delay (NCTSTD) system encountered in fed-batch production of 1,3-propanediol (1,3-PD). That is to say, we directly address the DRO problem governed by a CTD system, rather than discretizing the CTD system.

1,3-PD is a key chemical raw material for polymer production because of its unique symmetric structure [13]. The methods for the production of 1,3-PD can be categorized into two types: microbial conversion and chemical synthesis [14]. This paper focuses on the first category due to its eco-friendliness [15]. 1,3-PD microbial conversion process of glycerol can be divided into three categories [16]: (i) batch culture [17], (ii) continuous culture [18, 19], and (iii) fed-batch culture (see Figure 1 in [20]). In fed-batch culture, substrate inhibition is dramatically decreased, allowing for greater consumption of glycerol and alkali, which in turn leads to increased biomass production and higher concentrations of 1,3-PD. For fed-batch culture, many interesting papers appeared in the existing literature such as those involving state-dependent impulsive system [21], state-dependent switched system [22], multistage optimal control [23], optimal control [24], a novel downstream process [25], and Koopman modeling [26, 27]. However, the references mentioned do not consider the presence of time delays.

Time delays widely exist in many practical problems, including optimal control problem [28], predator-prey model [29], perimeter control problem [30], the survey [31], sequential time scaling transformation [32], non-autonomous time-delayed SIR model [33], multiregion urban traffic network [34], and sparse optimal control [35]. At the beginning of the fed-batch process, all substrates are placed into the fermentator. However, the chemical reaction does not trigger immediately [36], but needs to go through the process in which all substrates need to be well-mixed. Clearly, completing this process will require a certain amount of time. Thus, this process should be characterized by time-delay

systems. In the context of fed-batch culture with time-delay, numerous intriguing papers have been published in the existing literature, such as robust parameter identification [37], multi-objective optimization [38], and optimal control [39]. However, the aforementioned papers assumed that the probability distribution of the output measurement data is fully available. This assumption is obviously not realistic. In practical applications, typically only first-order moment information is obtainable. This is the main motivation for this paper.

This paper examines an NCTSTD system with unknown time-delay and system parameters to characterize the bioconversion process of 1,3-PD in fed-batch culture of glycerol caused by *K. pneumoniae*. Furthermore, the uncertain measurement output data is regarded as a stochastic variable and the only first-order moment information of its distribution is obtainable. Our aim is to identify these unknown quantities in a setting with uncertain output data measurement. Considering these factors, we propose a distributionally robust parameter estimation (DRPE) problem governed by the NCTSTD system, where the objective function contains two levels: (i) the inner-level is to maximize the expectation of the discrepancy between the output of the NCTSTD system and the uncertain output data measurement, and (ii) the outer-level is to minimize with respect to the unknown quantities. By the duality theory, the DRPE problem is equivalently translated into a single-level parameter estimation (SLPE) problem with non-smooth term smoothed by a smoothing function. Its convergence analysis is also given. Moreover, the gradients of the objective function are derived. A hybrid algorithm is proposed for solving the SLPE problem. Finally, numerical results show the applicability of the NCTSTD system to characterize the process of fed-batch culture and the availability of the proposed hybrid algorithm for solving the SLPE problem.

The contributions of the paper are threefold:

1) Different from the above literature, the optimal solution of the DRPE problem (OSDRPEP) is between the optimal solution of the robust parameter estimation problem (OSRPEP) and the optimal solution of the stochastic parameter estimation problem (OSSPEP). OSDRPEP builds a bridge between OSRPEP and OSSPEP, neither as conservative as OSRPEP nor as strict as OSSPEP when dealing with random probability distribution.

2) The DRPE problem subject to the CTD system is actually a bi-level parameter estimation problem, which cannot be described as an SLPE problem with finitely many equality constraints as the example of the DTD system. In this paper, a method on the basis of the duality theory is proposed for solving the DRPE problem governed by the CTD system.

3) The non-smooth term in the objective function is approximated by a smoothing technique. Some noteworthy theorems are required to prove that the smoothing function can overcome the problem that the constraint qualification is not satisfied.

The remainder of the paper is structured as follows. A DRPE problem is proposed in Section 2. A computational method is proposed for solving the DRPE problem in Section 3. Numerical results are given in Section 4. In Section 5, we draw some concluding remarks and imply some research directions in the future.

2. Problem formulation

2.1. NCTSTD system

In fed-batch culture, we assume that the conditions given below are satisfied.

- The concentrations of reactants are uniformly distributed in the reactor. Nonuniform space distribution is ignored.
- In fed-batch culture, the reactor is fed exclusively with glycerol and alkali.

Table 1. Nomenclature

Symbol	Representation	Symbol	Representation
$\mathbb{R}$	The set of real numbers	$\mathbb{R}_+$	The set of nonnegative real numbers
$\mathbb{N}_+$	The set of positive integers	t_f	The terminal time
Λ_1	$\{0, 1, 2, ..., N-1\}, N \in \mathbb{N}_+$	Λ_2	$\{0, 1, 2, ..., N\}, N \in \mathbb{N}_+$
$x_1(t)$	The concentration of biomass at time $t \in [0, t_f]$	$x_2(t)$	The concentration of glycerol at time $t \in [0, t_f]$
$x_3(t)$	The concentration of 1,3-PD at time $t \in [0, t_f]$	$x_4(t)$	The concentration of acetate at time $t \in [0, t_f]$
$x_5(t)$	The concentration of ethanol at time $t \in [0, t_f]$	$\phi(\cdot)$	A given initial function
x_0	The initial state	I_n	The set $\{1, 2, \ldots, n\}, n \in \mathbb{N}_+$
A^{T}	The transpose of the vector or matrix A	r	The velocity ratio of feeding alkali and glycerol
$q_1(t)$	The specific growth rate of cells at time $t \in [0, t_f]$	$D(t)$	The dilution rate at time $t \in [0, t_f]$
$q_2(t)$	The specific consumption rate of substrate at time $t \in [0, t_f]$	$q_\ell(t), \ell \in \{3, 4, 5\}$	The specific formation rates of product at time $t \in [0, t_f]$
$\tau^{2i+1}, i \in \Lambda_1$	The moments at each of which glycerol is added and the culture process switches from batch process to continuous process decided a priori in the experiment	$\tau^{2i+2}, i \in \Lambda_1$	The moments at each of which the flow of glycerol is ended and the culture process jumps from continuous process into batch process decided a priori in the experiment
C_{s_0}	The concentration of initial feed of substrate in the medium	V_0	The initial volume of solution in the fermentor
k_1	The Monod saturation constant for the substrate (mmol L^{-1})	p_2	The maintenance term of substrate consumption (mmol $g^{-1}h^{-1}$)
p_3	The maintenance term of 1,3-PD formation (mmol $g^{-1}h^{-1}$)	m_4	The maintenance term of acetate formation (mmol $g^{-1}h^{-1}$)
p_4	The maximum biomass growth yield in (g $mmol^{-1}$	p_5	The maximum extracellular 1,3-PD yield in (mmol g^{-1})
Y_4	The maximum acetate yield in (mmol g^{-1})	k_2	The saturation constants of substrate in the kinetic equations with excess terms (mmol L^{-1})
c_1, c_1, p_8, p_9	The parameters for the determination of yield of ethanol on glycerol	k_3	The saturation constants of extracellular 1,3-PD in the kinetic equations with excess terms (mmol L^{-1})
k_4	The saturation constants of EtOH in the kinetic equations with excess terms (mmol L^{-1})	p_1	The maximum specific growth rate (h^{-1})
p_6	The maximum increment of substrate consumption rate (h^{-1})	p_7, Δ_4	The maximum increments of product formation rates (h^{-1})
$\nu_\ell \geq 0$	The feeding rate of glycerol in $(\tau^{\ell-1}, \tau^\ell], \ell \in I_{2N+1}$	$\underline{p}_j$	The lower bound of $p_j, j \in I_9$
$\overline{p}_j$	The upper bound of $p_j, j \in I_9$	$\underline{x}_j$	The lower bound of $x_j(t), t \in [0, t_f], j \in I_5$
$\overline{x}_j$	The upper bound of $x_j(t), t \in [0, t_f], j \in I_5$	$\overline{h}$	The upper bound of h
h	The time-delay to be optimizated		

Based on [43], the NCTSTD system governing the fed-batch culture can be described as follows:

$$\begin{cases} \dfrac{\mathrm{d}x(t)}{\mathrm{d}t} = f^\ell(x(t), x(t-h), p) \in \mathbb{R}^5, \\ x(\tau^{\ell-1}+) = x(\tau^{\ell-1}), t \in (\tau^{\ell-1}, \tau^\ell], \ell \in I_{2N+1}, \\ x(0) = x_0, \\ x(t) = \phi(t), t \leq 0. \end{cases} \tag{2.1}$$

For $t \in (\tau^{\ell-1}, \tau^{\ell}], \ell = 2i+1, i \in \Lambda_2$, $f^{\ell}(x(t), x(t-h), p)$ is of the form given by

$$f^{2i+1}(x(t), x(t-h), p) := [f_1^{2i+1}(x(t), x(t-h), p), \ldots, f_5^{2i+1}(x(t), x(t-h), p)]^{\mathrm{T}},$$

where its components are

$$\begin{cases} f_j^{2i+1}(x(t), x(t-h), p) = q_j(t)x_1(t-h), j \in \{1,3,4,5\}, \\ f_2^{2i+1}(x(t), x(t-h), p) = -q_2(t)x_1(t-h). \end{cases} \tag{2.2}$$

On the other hand, for $t \in (\tau^{\ell-1}, \tau^{\ell}], \ell = 2i+2, i \in \Lambda_1$, $f^{\ell}(x(t), x(t-h), p)$ is of the form given by

$$f^{2i+2}(x(t), x(t-h), p) := [f_1^{2i+2}(x(t), x(t-h), p), \ldots, f_5^{2i+2}(x(t), x(t-h), p)]^{\mathrm{T}},$$

where its components are

$$\begin{cases} f_j^{2i+2}(x(t), x(t-h), p) = q_j(t)x_1(t-h) - D(t)x_j(t), j \in \{1,3,4,5\}, \\ f_2^{2i+2}(x(t), x(t-h), p) = D(t)\Big[\dfrac{C_{s_0}}{1+r} - x_2(t)\Big] - q_2(t)x_1(t-h), \end{cases} \tag{2.3}$$

with

$$D(t) = \frac{(1+r)\nu_{\ell}}{V(t)}, \tag{2.4}$$

$$V(t) = V_0 + \sum_{j=1}^{\ell-1}(1+r)\nu_j(\tau^j - \tau^{j-1}) + (1+r)\nu_{\ell}(t - \tau^{\ell-1}), \tag{2.5}$$

$$q_1(t) = \frac{p_1 x_2(t)}{x_2(t) + k_1}\prod_{j=2}^{5}(1 - \frac{x_j(t)}{x_j^*}), \tag{2.6}$$

$$q_2(t) = p_2 + q_1(t)p_4 + \frac{p_6 x_2(t)}{x_2(t) + k_2}, \tag{2.7}$$

$$q_3(t) = p_3 + q_1(t)p_5 + \frac{p_7 x_2(t)}{x_2(t) + k_3}, \tag{2.8}$$

$$q_4(t) = m_4 + q_1(t)Y_4 + \frac{\Delta_4 x_2(t)}{x_2(t) + k_4}, \tag{2.9}$$

$$q_5(t) = q_2(t)\Big[\frac{c_1}{c_2 + q_1(t)x_2(t)} + \frac{p_8}{p_9 + q_1(t)x_2(t)}\Big]. \tag{2.10}$$

The concentrations of substrate and the product of fermentation fluid are restricted to within a set W defined by

$$x(t) \in W := \prod_{j=1}^{5}[\underline{x}_j, \overline{x}_j] \subset \mathbb{R}_+^5, \forall t \in [0, t_f]. \tag{2.11}$$

Based on sensitivity analysis in [43], the time-delay h and the kinetic parameter vector p, the meanings of which are listed in Table 1, are required to be identified. Such analysis allows us to focus our

optimization efforts on these parameters and reduce unnecessary complexity. Thus, they are treated as decision variables to be chosen from the sets defined by

$$h \in [0, \overline{h}], \quad p := (p_1, \cdots, p_9)^{\mathrm{T}} \in \mathcal{P} := \prod_{\iota=1}^{9} [\underline{p}_\iota, \overline{p}_\iota].$$

This prior information helps narrow down the parameter space, making the optimization process more stable.

2.2. Distributionally robust parameter estimation problem

The aim of this section is to present a DRPE problem subject to the NCTSTD system with the unknown time-delay h and system parameter vector p to be optimized.

Let $x(\cdot|h,p) := (x_1(\cdot|h,p), ..., x_5(\cdot|h,p))^{\mathrm{T}}$ be the solution of system (2.1) (i.e., the system output) corresponding to each pair $(h,p) \in [0,\bar{h}] \times \mathcal{P}$. In [20, 37–39], assume that the distributions of output data are known precisely is an idealistic assumption. In fact, the distributions of the measured output data are often not available. Thus, the measured output data vector $\tilde{d}$ is regarded as a random vector. Its first-order moment is defined by

$$\mathbf{E}_{\mathbb{P}}[\tilde{d}] = \tilde{z}, \tag{2.12}$$

where $\mathbf{E}_{\mathbb{P}}$ represents the expected value according to the probability distribution $\mathbb{P}$; $\tilde{d} := ((\tilde{d}^1)^{\mathrm{T}}, \cdots, (\tilde{d}^\sigma)^{\mathrm{T}})^{\mathrm{T}} = (\tilde{d}_1^1, \cdots, \tilde{d}_3^1, \cdots, \tilde{d}_1^\sigma, \cdots, \tilde{d}_3^\sigma)^{\mathrm{T}} \in \mathbb{R}^{3\sigma}$; $\tilde{d}^\upsilon$ refers to the output vector recorded at specific sample times $t_\upsilon, \upsilon \in I_\sigma$ subject to $0 \leqslant t_1 < \cdots < t_\sigma \leqslant t_f$; and $\tilde{z} := (\tilde{z}_1^1, \cdots, \tilde{z}_3^1, \cdots, \tilde{z}_1^\sigma, \cdots, \tilde{z}_3^\sigma)^{\mathrm{T}} \in \mathbb{R}^{3\sigma}$.

The support set $\widetilde{\mathcal{D}}$ of $\tilde{d}$ is defined as

$$\widetilde{\mathcal{D}} = \left\{\tilde{d} \in \mathbb{R}^{3\sigma} : \underline{d}_\iota^\upsilon \leqslant \tilde{d}_\iota^\upsilon \leqslant \overline{d}_\iota^\upsilon, \upsilon \in I_\sigma, \iota \in I_3\right\}, \tag{2.13}$$

where $\underline{d}_\iota^\upsilon$ and $\overline{d}_\iota^\upsilon$ are, respectively, the lower and the upper limits of $\tilde{d}_\iota^\upsilon$ with $\underline{d}_\iota^\upsilon < \overline{d}_\iota^\upsilon$.

The ambiguity set $\widetilde{\mathfrak{P}}$ of probability measures of $\tilde{d}$ is defined as

$$\widetilde{\mathfrak{P}} = \left\{\mathbb{P} : \mathbb{P}(\tilde{d} \in \widetilde{\mathcal{D}}) = 1, \mathbf{E}_{\mathbb{P}}[\tilde{d}] = \tilde{z}\right\}. \tag{2.14}$$

Definition 1. *For a given $(h,p) \in [0,\bar{h}] \times \mathcal{P}$, the discrepancy between the system output $x_\iota(t_\upsilon|h,p), t_\upsilon \in \mathcal{T}, \iota \in I_3, \upsilon \in I_\sigma$, and the measured output data $\tilde{d}_\iota^\upsilon, \iota \in I_3, \upsilon \in I_\sigma$, is defined by*

$$R_e(h,p,\tilde{d}) := \frac{1}{3\sigma} \sum_{\iota=1}^{3} \sum_{\upsilon=1}^{\sigma} \left\{\frac{x_\iota(t_\upsilon|h,p) - \tilde{d}_\iota^\upsilon}{\overline{d}_\iota^\upsilon}\right\}^2, \tag{2.15}$$

where $\mathcal{T}$ represents the set of specified time points in fed-batch culture, and $|\mathcal{T}| = \sigma$ represents the number of elements in the set $\mathcal{T}$.

It is assumed that $\mathbb{P} \in \widetilde{\mathfrak{P}}$. Therefore, the worst-case scenario for $R_e(h,p,\tilde{d})$ is defined as $\max_{\mathbb{P}\in\widetilde{\mathfrak{P}}} \mathbf{E}_{\mathbb{P}}[R_e(h,p,\tilde{d})]$. Considering this, our DRPE problem is stated below:

$$\text{Problem } \mathbf{A}: \quad \min_{(h,p)\in[0,\bar{h}]\times\mathcal{P}} \left\{\max_{\mathbb{P}\in\widetilde{\mathfrak{P}}} \mathbf{E}_{\mathbb{P}}[R_e(h,p,\tilde{d})]\right\}$$

$$
\begin{aligned}
s.t. \quad & \mathbb{P}(\tilde{d} \in \widetilde{\mathcal{D}}) = 1, \\
& \mathbf{E}_{\mathbb{P}}[\tilde{d}] = \tilde{z}, \\
& g_j(x(t|h,p)) \leqslant 0, \forall t \in [0, t_f], j \in I_{10},
\end{aligned}
$$

where $g_j(x(t|h,p)) = -x_j(t|h,p), g_{j+5}(x(t|h,p)) = x_j(t|h,p) - \bar{x}_j, j \in I_5$.

3. Solution approaches

This section presents a numerical approach for solving Problem **A**. Figure 1 summarizes the approximation and transformation process of solving Problem **A**. A thorough discussion is given in Section 3.1. Based on Theorems 3 and 4, the aim of Theorem 3 is to calculate the gradients of $J^{\mathbf{H}}_{\rho,\epsilon}(h,p,\psi,\zeta)$ in Problem **H**. Theorems 1 and 2 are applied to carry out the error analysis. Algorithm 1 is employed to solve Problem **H**.

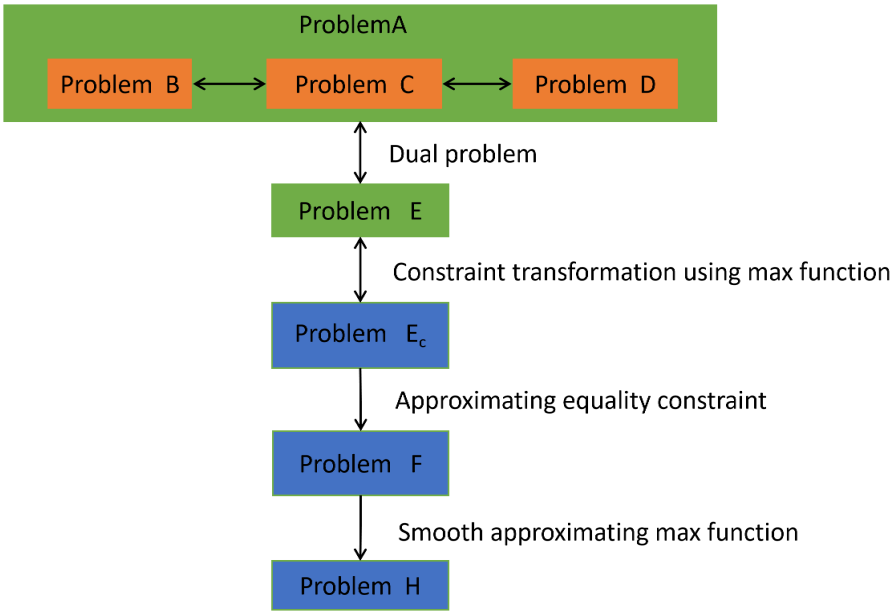

Figure 1. Solution process for Problem **A**.

3.1. Single-level parameter estimation problem

Problem **A** is a bi-level parameter estimation problem subject to the NCTSTD system. In this subsection, Problem **A** is converted into a deterministic single-level parameter estimation problem.

The inner-level problem of Problem **A** can be expressed as given below.

$$
\begin{aligned}
\text{Problem } \mathbf{B}: \quad \max_{\mathbb{P} \in \widetilde{\mathfrak{P}}} \quad & \mathbf{E}_{\mathbb{P}}[R_e(h,p,\tilde{d})] \\
s.t. \quad & \mathbb{P}(\tilde{d} \in \widetilde{\mathcal{D}}) = 1, \\
& \mathbf{E}_{\mathbb{P}}[\tilde{d}] = \tilde{z}.
\end{aligned}
$$

For ease of use, the expectation of the measured output data vector $\tilde{d}$ is expressed using the Lebesgue integral. Then, Problem **B** is reformulated as:

$$\begin{aligned} \text{Problem } \mathbf{C}: \quad \max_{\mathbb{P}\in\widetilde{\mathfrak{P}}} \quad & \int_{\widetilde{\mathcal{D}}} R_e(h,p,\tilde{d})\mathrm{d}\mathbb{P} \\ s.t. \quad & \int_{\widetilde{\mathcal{D}}} \mathbf{1}_{(\tilde{d}\in\widetilde{\mathcal{D}})}\mathrm{d}\mathbb{P}=1, \\ & \int_{\widetilde{\mathcal{D}}} \tilde{d}\mathrm{d}\mathbb{P}=\tilde{z}. \end{aligned}$$

Problem **C** is difficult to solve since it is a parameter estimation problem with respect to $\mathbb{P}\in\widetilde{\mathfrak{P}}$ on $\widetilde{\mathcal{D}}$. Considering this, in view of the duality theory [42], Problem **C** can be converted into its dual problem (i.e., Problem **D**), which can be solved easier than Problem **C**. Through the application of the duality theory [42], we obtain the duality of Problem **C** as follows:

$$\begin{aligned} \text{Problem } \mathbf{D}: \quad \min_{\zeta\in\mathbb{R}^{3\sigma},\psi\in\mathbb{R}} \quad & \psi+\tilde{z}^{\mathrm{T}}\zeta \\ s.t. \quad & \psi+d^{\mathrm{T}}\zeta\geqslant R_e(h,p,\tilde{d}),\forall d\in\widetilde{\mathcal{D}}, \end{aligned} \tag{3.1}$$

where $\zeta:=(\zeta_1^1,\cdots,\zeta_3^1,\cdots,\zeta_1^\sigma,\cdots,\zeta_3^\sigma)^{\mathrm{T}}\in\mathbb{R}^{3\sigma}$, ψ is the dual variable, and $R_e(h,p,\tilde{d})$ is as defined in (2.15).

Based on the definition of $R_e(h,p,\tilde{d})$ in (2.15) and max, we have

$$\begin{aligned} (3.1) \iff & \sum_{\iota=1}^{3}\sum_{\upsilon=1}^{\sigma}(\frac{d_\iota^\upsilon}{\overline{d}_\iota^\upsilon})^2-\sum_{\iota=1}^{3}\sum_{\upsilon=1}^{\sigma}(2\frac{x_\iota(t_\upsilon|h,p)}{\overline{d}_\iota^\upsilon}+\zeta_\iota^\upsilon)\frac{d_\iota^\upsilon}{\overline{d}_\iota^\upsilon}-\psi+\sum_{\iota=1}^{3}\sum_{\upsilon=1}^{\sigma}(\frac{x_\iota(t_\upsilon|h,p)}{\overline{d}_\iota^\upsilon})^2\leqslant 0,\forall d\in\widetilde{\mathcal{D}} \\ \iff & \max_{d\in\widetilde{\mathcal{D}}}\left\{\sum_{\iota=1}^{3}\sum_{\upsilon=1}^{\sigma}\left[(\frac{d_\iota^\upsilon}{\overline{d}_\iota^\upsilon})^2-(2\frac{x_\iota(t_\upsilon|h,p)}{\overline{d}_\iota^\upsilon}+\zeta_\iota^\upsilon)\frac{d_\iota^\upsilon}{\overline{d}_\iota^\upsilon}+(\frac{x_\iota(t_\upsilon|h,p)}{\overline{d}_\iota^\upsilon})^2\right]\right\}-\psi\leqslant 0. \end{aligned} \tag{3.2}$$

Based on the separability of the variables $d_\iota^\upsilon,\iota\in I_3,\upsilon\in I_\sigma$, constraint (3.2) is equivalent to constraint (3.3).

$$(3.2) \iff \sum_{\iota=1}^{3}\sum_{\upsilon=1}^{\sigma}\max_{d_\iota^\upsilon\in[\underline{d}_\iota^\upsilon,\overline{d}_\iota^\upsilon]}\left\{\left[(\frac{d_\iota^\upsilon}{\overline{d}_\iota^\upsilon})^2-(2\frac{x_\iota(t_\upsilon|h,p)}{\overline{d}_\iota^\upsilon}+\zeta_\iota^\upsilon)\frac{d_\iota^\upsilon}{\overline{d}_\iota^\upsilon}+(\frac{x_\iota(t_\upsilon|h,p)}{\overline{d}_\iota^\upsilon})^2\right]\right\}-\psi\leqslant 0. \tag{3.3}$$

From the properties of quadratic functions, it follows that constraint (3.3) $\Leftrightarrow$ constraint (3.4).

$$\begin{aligned} (3.3) \iff & \sum_{\iota=1}^{3}\sum_{\upsilon=1}^{\sigma}\max\left\{(\frac{d_\iota^\upsilon}{\overline{d}_\iota^\upsilon})^2-(2\frac{x_\iota(t_\upsilon|h,p)}{\overline{d}_\iota^\upsilon}+\zeta_\iota^\upsilon)\frac{d_\iota^\upsilon}{\overline{d}_\iota^\upsilon}+(\frac{x_\iota(t_\upsilon|h,p)}{\overline{d}_\iota^\upsilon})^2,\right. \\ & \left.1-(2\frac{x_\iota(t_\upsilon|h,p)}{\overline{d}_\iota^\upsilon}+\zeta_\iota^\upsilon)+(\frac{x_\iota(t_\upsilon|h,p)}{\overline{d}_\iota^\upsilon})^2\right\}-\psi\leqslant 0. \end{aligned} \tag{3.4}$$

Since $\max\left\{a_1,a_2\right\}=\dfrac{a_1+a_2}{2}+\dfrac{|a_1-a_2|}{2}$, constraint (3.4) is equivalent to constraint (3.5).

$$(3.4) \iff \frac{1}{2}\sum_{\iota=1}^{3}\sum_{\upsilon=1}^{\sigma}\left\{(\frac{d_\iota^\upsilon}{\overline{d}_\iota^\upsilon})^2+1-(2\frac{x_\iota(t_\upsilon|h,p)}{\overline{d}_\iota^\upsilon}+\zeta_\iota^\upsilon)(\frac{d_\iota^\upsilon}{\overline{d}_\iota^\upsilon}+1)+2(\frac{x_\iota(t_\upsilon|h,p)}{\overline{d}_\iota^\upsilon})^2+\right.$$

$$\left|(\frac{\underline{d}_\iota^\upsilon}{\overline{d}_\iota^\upsilon}-1)(\frac{\underline{d}_\iota^\upsilon}{\overline{d}_\iota^\upsilon}+1-2\frac{x_\iota(t_\upsilon|h,p)}{\overline{d}_\iota^\upsilon}-\zeta_\iota^\upsilon)\right|\Big\}-\psi\leqslant 0. \tag{3.5}$$

Letting $\alpha_\iota^\upsilon(x(t_\upsilon|h,p),\zeta):=(\frac{\underline{d}_\iota^\upsilon}{\overline{d}_\iota^\upsilon})^2+1-(2\frac{x_\iota(t_\upsilon|h,p)}{\overline{d}_\iota^\upsilon}+\zeta_\iota^\upsilon)(\frac{\underline{d}_\iota^\upsilon}{\overline{d}_\iota^\upsilon}+1)+2(\frac{x_\iota(t_\upsilon|h,p)}{\overline{d}_\iota^\upsilon})^2;\beta_\iota^\upsilon(x(t_\upsilon|h,p),\zeta):=(\frac{\underline{d}_\iota^\upsilon}{\overline{d}_\iota^\upsilon}-1)(\frac{\underline{d}_\iota^\upsilon}{\overline{d}_\iota^\upsilon}+1-2\frac{x_\iota(t_\upsilon|h,p)}{\overline{d}_\iota^\upsilon}-\zeta_\iota^\upsilon)$. Based on $\max\left\{a_1,a_2\right\}=\frac{a_1+a_2}{2}+\frac{|a_1-a_2|}{2}$, we have

$$\begin{aligned}(3.5)\iff&\frac{1}{2}\sum_{\iota=1}^{3}\sum_{\upsilon=1}^{\sigma}\left\{\alpha_\iota^\upsilon(x(t_\upsilon|h,p),\zeta)+2\max\left\{\beta_\iota^\upsilon(x(t_\upsilon|h,p),\zeta),0\right\}-\beta_\iota^\upsilon(x(t_\upsilon|h,p),\zeta)\right\}-\psi\leqslant 0\\ \iff&\max\left\{\Lambda(h,p,\zeta,\psi),0\right\}=0,\end{aligned} \tag{3.6}$$

where $\Lambda(h,p,\zeta,\psi):=\frac{1}{2}\sum_{\iota=1}^{3}\sum_{\upsilon=1}^{\sigma}\left\{\alpha_\iota^\upsilon(x(t_\upsilon|h,p),\zeta)+2\max\left\{\beta_\iota^\upsilon(x(t_\upsilon|h,p),\zeta),0\right\}-\beta_\iota^\upsilon(x(t_\upsilon|h,p),\zeta)\right\}-\psi$.

Problem **A** is equivalently transformed into the following SLPE problem:

$$\begin{aligned}\text{Problem }\mathbf{E}:\ \min_{\zeta\in\mathbb{R}^{3\sigma},\psi\in\mathbb{R},h\in[0,\bar{h}],p\in\mathcal{P}}\ &J^{\mathbf{E}}(h,p,\psi,\zeta):=\psi+\tilde{z}^{\mathrm{T}}\zeta\\ s.t.\ \ &g_j(x(t|h,p))\leqslant 0,\forall t\in[0,t_f],j\in I_{10},\\ &\Lambda(h,p,\zeta,\psi)\leqslant 0.\end{aligned} \tag{3.7}$$

3.2. Constraint transcription technique

Problem **E** involves an infinite number of constraints (i.e., constraint (3.8) that restricts the state variables at each point in $[0,t_f]$). Therefore, Problem **E** can be regarded as a semi-infinite programming problem [46]. Based on the constraint transformation method involving the definition of $\max\{\cdot,\cdot\}$ [46], Problem **E** is equivalently transformed into Problem $\mathbf{E_c}$.

$$\begin{aligned}\text{Problem }\mathbf{E_c}:\ \min_{\zeta\in\mathbb{R}^{3\sigma},\psi\in\mathbb{R},h\in[0,\bar{h}],p\in\mathcal{P}}\ &J^{\mathbf{E}}(h,p,\psi,\zeta):=\psi+\tilde{z}^{\mathrm{T}}\zeta\\ s.t.\ \ &\sum_{j=1}^{10}\int_0^{t_f}\max\left\{g_j(x(t|h,p)),0\right\}=0,\\ &\max\left\{\Lambda(h,p,\zeta,\psi),0\right\}=0.\end{aligned} \tag{3.8}$$

In view of the penalty function method [45, 46], Problem $\mathbf{E_c}$ can be approximated by Problem **F**.

$$\begin{aligned}\text{Problem }\mathbf{F}:\min_{\zeta\in\mathbb{R}^{3\sigma},\psi\in\mathbb{R},h\in[0,\bar{h}],p\in\mathcal{P}}J_\rho^{\mathbf{F}}(h,p,\psi,\zeta)\ :=\ &J^{\mathbf{E}}(h,p,\psi,\zeta)+\rho\sum_{j=1}^{10}\int_0^{t_f}\max\left\{g_j(x(t|h,p)),0\right\}\mathrm{d}t\\ &+\rho\max\left\{\Lambda(h,p,\zeta,\psi),0\right\}.\end{aligned} \tag{3.9}$$

Remark 1. *It can be demonstrated that if the penalty parameter ρ is sufficiently increased, then any minimizer of Problem* **F** *within the region $\zeta\in\mathbb{R}^{3\sigma},\psi\in\mathbb{R},h\in[0,\bar{h}],p\in\mathcal{P}$ will be feasible for Problem* $\mathbf{E_c}$*. Therefore, a solution to Problem* $\mathbf{E_c}$ *can be achieved by minimizing Problem* **F** *with appropriately chosen values for the penalty parameter ρ.*

3.3. Smoothing technique and error analysis

3.3.1. Smoothing technique

One of the main hardships to solve Problem **F** numerically is that the second and the third terms in (3.9) are non-smooth, with respect to h and p, due to the existence of the function $\max\{\cdot, 0\}$. To surmount this hardship, the function $\max\{\cdot, 0\}$ is approximated by the smoothing technique [44] as follows.

$$\gamma_{\epsilon,\rho,\kappa}[\eta] = \begin{cases} 0, & \text{if } \eta \leq -\dfrac{\epsilon}{\kappa\rho}, \\ \dfrac{\kappa\rho}{2\epsilon}\eta^2 + \eta + \dfrac{\epsilon}{2\kappa\rho}, & \text{if } -\dfrac{\epsilon}{\kappa\rho} \leq \eta \leq 0, \\ \eta + \dfrac{\epsilon}{2\kappa\rho}, & \text{if } \eta > 0, \end{cases} \tag{3.10}$$

where η is a continuous inequality constraint; ρ is a penalty factor, κ is the number of continuous inequality constraints, and $\epsilon > 0$ is called the smoothing parameter.

Approximate the non-smooth terms in Problem **F** using the function (3.10), and we get

$$\begin{aligned} \text{Problem } \mathbf{H}: \min_{\zeta\in\mathbb{R}^{3\sigma},\psi\in\mathbb{R},h\in[0,\bar{h}],p\in\mathcal{P}} J^{\mathbf{H}}_{\rho,\epsilon}(h,p,\psi,\zeta) \;:=\; & J^{\mathbf{E}}(h,p,\psi,\zeta) + \rho\sum_{j=1}^{10}\int_0^{t_f}\gamma_{\epsilon,\rho,10}[g_j(x(t|h,p))]\mathrm{d}t \\ & +\rho\gamma_{\epsilon,\rho,1}[\Lambda_\epsilon(h,p,\zeta,\psi)], \end{aligned}$$

where $\Lambda_\epsilon(h,p,\zeta,\psi) := \dfrac{1}{2}\sum_{\iota=1}^{3}\sum_{\upsilon=1}^{\sigma}\left\{\alpha_\iota^\upsilon(x(t_\upsilon|h,p),\zeta) + 2\gamma_{\epsilon,\rho,3\sigma}[\beta_\iota^\upsilon(x(t_\upsilon|h,p),\zeta)] - \beta_\iota^\upsilon(x(t_\upsilon|h,p),\zeta)\right\} - \psi.$

3.3.2. Error analysis

Since the functions involved with Problem **H** are smooth, it can be addressed using the gradient-based nonlinear programming (NLP) algorithm. It is evident that the difference between the solution of Problem **H** and the solution of Problem **E** is of importance. This section conducts an error analysis comparing Problem **H**, Problem **F**, and Problem **E**. With this in mind, Lemmas 1 and 2 are needed in the process of error analysis.

Lemma 1. *For any $\epsilon > 0$, the function $\gamma_{\epsilon,\rho,\kappa}[\eta]$ has the following properties:*

$$0 \leqslant \gamma_{\epsilon,\rho,10}[g_j(x(t|h,p))] - \max\left\{g_j(x(t|h,p)), 0\right\} \leqslant \frac{\epsilon}{20\rho}, \tag{3.11}$$

$$0 \leqslant \gamma_{\epsilon,\rho,3\sigma}[\beta_\iota^\upsilon(x(t_\upsilon|h,p),\zeta)] - \max\left\{\beta_\iota^\upsilon(x(t_\upsilon|h,p),\zeta), 0\right\} \leqslant \frac{\epsilon}{6\sigma\rho}, \tag{3.12}$$

$$0 \leqslant \gamma_{\epsilon,\rho,1}[\Lambda_\epsilon(h,p,\zeta,\psi)] - \max\left\{\Lambda(h,p,\zeta,\psi), 0\right\} \leqslant \frac{\epsilon}{\rho}. \tag{3.13}$$

Proof. In view of Theorem 3.2.1 in [44], it is easy to prove the validity of (3.11) and (3.12) and

$$0 \leqslant \gamma_{\epsilon,\rho,1}[\Lambda_\epsilon(h,p,\zeta,\psi)] - \max\left\{\Lambda_\epsilon(h,p,\zeta,\psi), 0\right\} \leqslant \frac{\epsilon}{2\rho}. \tag{3.14}$$

Now, we will prove (3.13). Based on (3.12), we have

$$0 \;\leqslant\; \Lambda_\epsilon(h,p,\zeta,\psi) - \Lambda(h,p,\zeta,\psi) = \sum_{\iota=1}^{3}\sum_{\upsilon=1}^{\sigma}\left\{\gamma_{\epsilon,\rho,3\sigma}[\beta(x_\iota(t_\upsilon|h,p),\zeta_\iota^\upsilon)] - \max\left\{\beta(x_\iota(t_\upsilon|h,p),\zeta_\iota^\upsilon), 0\right\}\right\}$$

$$\leqslant \frac{\epsilon}{2\rho}, \tag{3.15}$$

which implies

$$0 \leqslant \max\left\{\Lambda_\epsilon(h,p,\zeta,\psi),0\right\} - \max\left\{\Lambda(h,p,\zeta,\psi),0\right\} \leqslant \Lambda_\epsilon(h,p,\zeta,\psi) - \Lambda(h,p,\zeta,\psi) \leqslant \frac{\epsilon}{2\rho}. \tag{3.16}$$

According to (3.14) and (3.16), it follows that

$$\begin{aligned} 0 &\leqslant \gamma_{\epsilon,\rho,1}[\Lambda_\epsilon(h,p,\zeta,\psi)] - \max\left\{\Lambda(h,p,\zeta,\psi),0\right\} \\ &= \gamma_{\epsilon,\rho,1}[\Lambda_\epsilon(h,p,\zeta,\psi)] - \max\left\{\Lambda_\epsilon(h,p,\zeta,\psi),0\right\} + \max\left\{\Lambda_\epsilon(h,p,\zeta,\psi),0\right\} - \max\left\{\Lambda(h,p,\zeta,\psi),0\right\} \\ &\leqslant \frac{\epsilon}{\rho}, \end{aligned} \tag{3.17}$$

which completes the proof. ■ □

Lemma 2. *For $\rho > 0$ and $\epsilon > 0$, it holds that*

$$0 \leqslant J^{\mathbf{H}}_{\rho,\epsilon}(h,p,\psi,\zeta) - J^{\mathbf{F}}_{\rho}(h,p,\psi,\zeta) \leqslant \epsilon(1+\frac{t_f}{2}).$$

Proof. Based on Lemma 1, we have

$$\begin{aligned} 0 &\leqslant J^{\mathbf{H}}_{\rho,\epsilon}(h,p,\psi,\zeta) - J^{\mathbf{F}}_{\rho}(h,p,\psi,\zeta) \\ &= \rho\sum_{j=1}^{10}\int_0^{t_f}\left\{\gamma_{\epsilon,\rho,10}[g_j(x(t|h,p))] - \max\left\{g_j(x(t|h,p)),0\right\}\right\}\mathrm{d}t \\ &\quad +\rho\left\{\gamma_{\epsilon,\rho,1}[\Lambda_\epsilon(h,p,\zeta,\psi)] - \max\left\{\Lambda(h,p,\zeta,\psi),0\right\}\right\} \\ &\leqslant \epsilon(1+\frac{t_f}{2}). \end{aligned} \tag{3.18}$$

The proof is complete. ■ □

Using Lemmas 1 and 2 as a foundation, we now proceed with the error analysis comparing Problem **H**, Problem **F**, and Problem **E**.

Theorem 1. *Let (h^*,p^*,ψ^*,ζ^*) and $(h^{**},p^{**},\zeta^{**},\psi^{**})$ be the solutions of Problems* **F** *and* **H**, *respectively. Then,*

$$0 \leqslant J^{\mathbf{H}}_{\rho,\epsilon}(h^{**},p^{**},\zeta^{**},\psi^{**}) - J^{\mathbf{F}}_{\rho}(h^*,p^*,\psi^*,\zeta^*) \leqslant \epsilon(1+\frac{t_f}{2}).$$

Proof. In view of Lemma 2,

$$0 \leqslant J^{\mathbf{H}}_{\rho,\epsilon}(h^{**},p^{**},\zeta^{**},\psi^{**}) - J^{\mathbf{F}}_{\rho}(h^{**},p^{**},\zeta^{**},\psi^{**}) \leqslant \epsilon(1+\frac{t_f}{2}), \tag{3.19}$$

$$0 \leqslant J^{\mathbf{H}}_{\rho,\epsilon}(h^*,p^*,\psi^*,\zeta^*) - J^{\mathbf{F}}_{\rho}(h^*,p^*,\psi^*,\zeta^*) \leqslant \epsilon(1+\frac{t_f}{2}). \tag{3.20}$$

Because $(h^*, p^*, \psi^*, \zeta^*)$ is the solution of Problem **F**, we have

$$-J_\rho^{\mathbf{F}}(h^{**}, p^{**}, \zeta^{**}, \psi^{**}) \leqslant -J_\rho^{\mathbf{F}}(h^*, p^*, \psi^*, \zeta^*),$$

which is equivalent to

$$J_{\rho,\epsilon}^{\mathbf{H}}(h^{**}, p^{**}, \zeta^{**}, \psi^{**}) - J_\rho^{\mathbf{F}}(h^{**}, p^{**}, \zeta^{**}, \psi^{**}) \leqslant J_{\rho,\epsilon}^{\mathbf{H}}(h^{**}, p^{**}, \zeta^{**}, \psi^{**}) - J_\rho^{\mathbf{F}}(h^*, p^*, \zeta^*, \psi^*). \tag{3.21}$$

As $(\tau^{**}, p^{**}, \theta^{**}, \eta^{**})$ is the solution of Problem $\mathbf{SLI_2}$, it is clear that

$$J_{\rho,\epsilon}^{\mathbf{H}}(h^{**}, p^{**}, \zeta^{**}, \psi^{**}) \leqslant J_{\rho,\epsilon}^{\mathbf{H}}(h^*, p^*, \zeta^*, \psi^*),$$

which is equivalent to

$$J_{\rho,\epsilon}^{\mathbf{H}}(h^{**}, p^{**}, \zeta^{**}, \psi^{**}) - J_\rho^{\mathbf{F}}(h^*, p^*, \zeta^*, \psi^*) \leqslant J_{\rho,\epsilon}^{\mathbf{H}}(h^*, p^*, \zeta^*, \psi^*) - J_\rho^{\mathbf{F}}(h^*, p^*, \zeta^*, \psi^*). \tag{3.22}$$

Thus, it follows from (3.19)–(3.22) that

$$\begin{aligned} 0 &\leqslant J_{\rho,\epsilon}^{\mathbf{H}}(h^{**}, p^{**}, \zeta^{**}, \psi^{**}) - J_\rho^{\mathbf{F}}(h^{**}, p^{**}, \zeta^{**}, \psi^{**}) \\ &\leqslant J_{\rho,\epsilon}^{\mathbf{H}}(h^{**}, p^{**}, \zeta^{**}, \psi^{**}) - J_\rho^{\mathbf{F}}(h^*, p^*, \zeta^*, \psi^*) \\ &\leqslant J_{\rho,\epsilon}^{\mathbf{H}}(h^*, p^*, \zeta^*, \psi^*) - J_\rho^{\mathbf{F}}(h^*, p^*, \zeta^*, \psi^*) \\ &\leqslant \epsilon(1 + \frac{t_f}{2}). \end{aligned}$$

The proof is complete. ■ □

Definition 2. *If $g_j(x(t|h^\epsilon, p^\epsilon)) \leqslant \epsilon, j \in I_{10}; \Lambda(h^\epsilon, p^\epsilon, \psi^\epsilon, \zeta^\epsilon) \leqslant \epsilon$, then the set of decision variables $(h^\epsilon, p^\epsilon, \psi^\epsilon, \zeta^\epsilon)$ is called ϵ–feasible to Problem* **E**.

Theorem 2. *Let $(h^*, p^*, \psi^*, \zeta^*)$ and $(h^{**}, p^{**}, \zeta^{**}, \psi^{**})$ be the solution of Problems* **F** *and* **H**, *respectively. Furthermore, suppose that $(h^*, p^*, \psi^*, \zeta^*)$ is feasible to Problem* **E** *and that $(h^{**}, p^{**}, \zeta^{**}, \psi^{**})$ is ϵ–feasible to Problem* **E**. *Then,*

$$-\frac{\epsilon[20\rho t_f + t_f + 2\rho + 2]}{2} \leqslant J^{\mathbf{E}}(h^{**}, p^{**}, \psi^{**}, \zeta^{**}) - J^{\mathbf{E}}(h^*, p^*, \psi^*, \zeta^*) \leqslant \epsilon(1 + \frac{t_f}{2}).$$

Proof. Due to the feasibility of $(h^*, p^*, \zeta^*, \psi^*)$ in Problem **E**, we get

$$\max\left\{\Lambda(h^*, p^*, \zeta^*, \psi^*), 0\right\} = 0, \quad \max\left\{g_j(x(t|h^*, p^*)), 0\right\}\mathrm{d}t = 0, j \in I_{10}. \tag{3.23}$$

From the ϵ–feasibility of $(h^{**}, p^{**}, \zeta^{**}, \psi^{**})$ in Problem **E**, it follows that

$$g_j(x(t|h^{**}, p^{**})) \leqslant \epsilon, j \in I_{10}, \tag{3.24}$$

$$\Lambda(h^{**}, p^{**}, \zeta^{**}, \psi^{**}) \leqslant \epsilon. \tag{3.25}$$

Based on (3.15) and (3.25), we obtain

$$\Lambda_\epsilon(h^{**}, p^{**}, \zeta^{**}, \psi^{**}) \leqslant \Lambda(h^{**}, p^{**}, \zeta^{**}, \psi^{**}) + \frac{\epsilon}{2\rho} \leqslant \epsilon + \frac{\epsilon}{2\rho}. \tag{3.26}$$

From (3.24), (3.26), and the monotonicity of the $\gamma_{\epsilon,\rho,\kappa}[\eta]$ in (3.10), it follows that

$$0 \leqslant \gamma_{\epsilon,\rho,10}[g_j(x(t|h^{**}, p^{**}))] \leqslant \epsilon + \frac{\epsilon}{20\rho},\ j \in I_{10};\quad 0 \leqslant \gamma_{\epsilon,\rho,1}[\Lambda_\epsilon(h^{**}, p^{**}, \zeta^{**}, \psi^{**})] \leqslant \epsilon + \frac{\epsilon}{\rho},$$

which implies

$$0 \leqslant \rho \sum_{j=1}^{10} \int_0^{t_f} \gamma_{\epsilon,\rho,10}[g_j(x(t|h^{**}, p^{**}))]\mathrm{d}t + \rho\gamma_{\epsilon,\rho,1}[\Lambda_\epsilon(h^{**}, p^{**}, \zeta^{**}, \psi^{**})] \leqslant \frac{\epsilon[20\rho t_f + t_f + 2\rho + 2]}{2}. \tag{3.27}$$

This can be rearranged to yield

$$\begin{aligned} -\frac{\epsilon[20\rho t_f + t_f + 2\rho + 2]}{2} &\leqslant -\rho \sum_{j=1}^{10} \int_0^{t_f} \gamma_{\epsilon,\rho,10}[g_j(x(t|h^{**}, p^{**}))]\mathrm{d}t - \rho\gamma_{\epsilon,\rho,1}[\Lambda_\epsilon(h^{**}, p^{**}, \zeta^{**}, \psi^{**})] \\ &\leqslant 0. \end{aligned} \tag{3.28}$$

It is clear from Theorem 1 that

$$\begin{aligned} 0 \leqslant\ & J^{\mathbf{E}}(h^{**}, p^{**}, \psi^{**}, \zeta^{**}) + \rho \sum_{j=1}^{10} \int_0^{t_f} \gamma_{\epsilon,\rho,10}[g_j(x(t|h^{**}, p^{**}))]\mathrm{d}t + \rho\gamma_{\epsilon,\rho,1}[\Lambda_\epsilon(h^{**}, p^{**}, \psi^{**}, \zeta^{**})] \\ & - J^{\mathbf{E}}(h^*, p^*, \zeta^*, \psi^*) - \rho \sum_{j=1}^{10} \int_0^{t_f} \max\left\{g_j(x(t|h^*, p^*)), 0\right\}\mathrm{d}t - \rho \max\left\{\Lambda(h^*, p^*, \zeta^*, \psi^*), 0\right\} \\ \leqslant\ & \epsilon(1 + \frac{t_f}{2}). \end{aligned}$$

Combining (3.23) and (3.27) gives

$$-\frac{\epsilon[20\rho t_f + t_f + 2\rho + 2]}{2} \leqslant J^{\mathbf{E}}(h^{**}, p^{**}, \psi^{**}, \zeta^{**}) - J^{\mathbf{E}}(h^*, p^*, \psi^*, \zeta^*) \leqslant \epsilon(1 + \frac{t_f}{2}).$$

The proof is complete. ■ □

3.4. State variation

In the process of the numerical optimization, the gradient formulae of $J^{\mathbf{H}}_{\rho,\epsilon}(h, p, \psi, \zeta)$ of Problem **H** with respect to h, p, ψ, and ζ are required. With these gradient formulae, a gradient-based NLP algorithm can be employed for solving Problem **H**. Therefore, this section is dedicated to deriving the gradient formulas.

3.4.1. State variation with respect to time-delay h

The partial derivative of $x(t|\cdot, p)$ with respect to h is known as the state variation with respect to h.

Lemma 3. *For each pair $(h, p) \in [0, \bar{h}] \times \mathcal{P}$, it holds that*

$$\frac{\partial x(t|h, p)}{\partial h} = \varpi(t|h, p) = (\varpi_1(t|h, p), \ldots, \varpi_5(t|h, p))^{\mathrm{T}}, t \in (\tau^{\ell-1}, \tau^\ell], \ell \in I_{2N+1}, \tag{3.29}$$

where $\varpi(t|h,p) \in \mathbb{R}^5$ is the solution of the following variational system:

$$\begin{cases} \dot{\varpi}(t) = \dfrac{\partial f^\ell(x(t),x(t-h),p)}{\partial x(t)}\varpi(t) + \dfrac{\partial f^\ell(x(t),x(t-h),p)}{\partial x(t-h)}\varpi(t-h) + \dfrac{\partial f^\ell(x(t),x(t-h),p)}{\partial x(t-h)}\chi(t-h), \\ \quad t \in (\tau^{\ell-1},\tau^\ell], \ell \in I_{2N+1}, \\ \varpi(\tau^{\ell-1}+) = \varpi(\tau^\ell), \ell \in I_{2N+1}, \\ \varpi(t) = 0, t \leqslant 0, \\ \chi(t) = \begin{cases} \dfrac{d\phi(t)}{dt}, t \leqslant 0, \\ f^\ell(x(t),x(t-h),p), t \in (\tau^{\ell-1},\tau^\ell], \ell \in I_{2N+1}. \end{cases} \end{cases} \tag{3.30}$$

Proof. This theorem can be proven in a manner similar to that used for Theorem 5 in [43]. ■ □

3.4.2. The state variation with respect to $p_\iota, \iota \in I_9$

To solve Problem **H**, we also require the information of the state variation with respect to $p_\iota, \iota \in I_9$.

Lemma 4. *For each pair $(h,p) \in [0,\bar{h}] \times \mathcal{P}$, it holds that*

$$\frac{\partial x(t|h,p)}{\partial p_\iota} = \pi_\iota(t|h,p) = (\pi_\iota^1(t|h,p),\ldots,\pi_\iota^5(t|h,p))^{\mathrm{T}}, t \in (\tau^{\ell-1},\tau^\ell], \ell \in I_{2N+1}, \iota \in I_9, \tag{3.31}$$

where $\pi_\iota(t|h,p) \in \mathbb{R}^5, \iota \in I_9$, is the solution to the following variational system:

$$\begin{cases} \dot{\pi}_\iota(t) = \dfrac{\partial f^\ell(x(t),x(t-h),p)}{\partial x(t)}\pi_\iota(t) + \dfrac{\partial f^\ell(x(t),x(t-h),p)}{\partial x(t-h)}\pi_\iota(t-h) + \dfrac{\partial f^\ell(x(t),x(t-h),p)}{\partial p_\iota}, \\ \quad t \in (\tau^{\ell-1},\tau^\ell], \ell \in I_{2N+1}, \iota \in I_9, \\ \pi_\iota(\tau^{\ell-1}+) = \pi_\iota(\tau^\ell), \ell \in I_{2N+1}, \iota \in I_9, \\ \pi_\iota(t) = 0, t \leqslant 0, \iota \in I_9. \end{cases} \tag{3.32}$$

Proof. This theorem can be proven in a manner similar to that used for Theorem 5 in [43]. ■ □

3.5. Gradient formulas

The state variation with respect to h and $p_\iota, \iota \in I_9$ can be computed through solving system (2.1) together with systems (3.30) and (3.32) forward in time. Then, the gradient formulae of $J^{\mathbf{H}}_{\rho,\epsilon}(h,p,\psi,\zeta)$ of Problem **H** with respect to h, p, ψ, and ζ can be easily obtained. By Lemmas 3 and 4 and the chain rule, we have

Theorem 3. *The gradients of the cost function $J^{\mathbf{H}}_{\rho,\epsilon}(h,p,\psi,\zeta)$ with respect to $\lambda \in \left\{h; p_\iota, \iota \in I_9; \psi; \zeta_\iota^\upsilon, \iota \in I_3, \upsilon \in I_\sigma\right\}$ are given by*

$$\frac{\partial J^{\mathbf{H}}_{\rho,\epsilon}(h,p,\psi,\zeta)}{\partial \lambda} = \frac{\partial J^{\mathbf{E}}(h,p,\psi,\zeta)}{\partial \lambda} + \rho\Big\{\sum_{j=1}^{10}\int_0^{t_f}\frac{\partial \gamma_{\epsilon,\rho,10}[g_j(x(t|h,p))]}{\partial \lambda}dt + \frac{\partial \gamma_{\epsilon,\rho,1}[\Lambda_\epsilon(h,p,\zeta,\psi)]}{\partial \lambda}\Big\},$$

where

$$\frac{\partial J^{\mathbf{E}}(h,p,\psi,\zeta)}{\partial h}=0;\ \frac{\partial J^{\mathbf{E}}(h,p,\psi,\zeta)}{\partial p_\iota}=0,\iota\in I_9;\ \frac{\partial J^{\mathbf{E}}(h,p,\psi,\zeta)}{\partial \psi}=1;\ \frac{\partial J^{\mathbf{E}}(h,p,\psi,\zeta)}{\partial \zeta_\iota^\upsilon}=\tilde{z}_\iota^\upsilon,\iota\in I_3,\upsilon\in I_\sigma;$$

$$\frac{\partial \gamma_{\epsilon,\rho,10}[g_j(x(t|h,p))]}{\partial \psi}=0;\ \frac{\partial \gamma_{\epsilon,\rho,10}[g_j(x(t|h,p))]}{\partial \zeta_\iota^\upsilon}=0,\iota\in I_3,\upsilon\in I_\sigma;$$

$$\frac{\partial \gamma_{\epsilon,\rho,10}[g_j(x(t|h,p))]}{\partial h}=\frac{\partial \gamma_{\epsilon,\rho,10}[g_j(x(t|h,p))]}{\partial g_j(x(t|h,p))}\frac{\partial g_j(x(t|h,p))}{\partial x(t|h,p))}\varpi(t|h,p)),j\in I_{10};$$

$$\frac{\partial \gamma_{\epsilon,\rho,10}[g_j(x(t|h,p))]}{\partial p_\iota}=\frac{\partial \gamma_{\epsilon,\rho,10}[g_j(x(t|h,p))]}{\partial g_j(x(t|h,p))}\frac{\partial g_j(x(t|h,p))}{\partial x(t|h,p))}\pi_\iota(t|h,p)),j\in I_{10},\iota\in I_9;$$

$$\frac{\partial \gamma_{\epsilon,\rho,1}[\Lambda_\epsilon(h,p,\zeta,\psi)]}{\partial h}=\frac{\partial \gamma_{\epsilon,\rho,1}[\Lambda_\epsilon(h,p,\zeta,\psi)]}{\partial \Lambda_\epsilon(h,p,\zeta,\psi)}\frac{\partial \Lambda_\epsilon(h,p,\zeta,\psi)}{\partial h};$$

$$\frac{\partial \gamma_{\epsilon,\rho,1}[\Lambda_\epsilon(h,p,\zeta,\psi)]}{\partial p_\iota}=\frac{\partial \gamma_{\epsilon,\rho,1}[\Lambda_\epsilon(h,p,\zeta,\psi)]}{\partial \Lambda_\epsilon(h,p,\zeta,\psi)}\frac{\partial \Lambda_\epsilon(h,p,\zeta,\psi)}{\partial p_\iota},\iota\in I_9;$$

$$\frac{\partial \gamma_{\epsilon,\rho,1}[\Lambda_\epsilon(h,p,\zeta,\psi)]}{\partial \zeta_\iota^\upsilon}=\frac{\partial \gamma_{\epsilon,\rho,1}[\Lambda_\epsilon(h,p,\zeta,\psi)]}{\partial \Lambda_\epsilon(h,p,\zeta,\psi)}\frac{\partial \Lambda_\epsilon(h,p,\zeta,\psi)}{\partial \zeta_\iota^\upsilon},\iota\in I_3,\upsilon\in I_\sigma;$$

$$\frac{\partial \gamma_{\epsilon,\rho,1}[\Lambda_\epsilon(h,p,\zeta,\psi)]}{\partial \psi}=-\frac{\partial \gamma_{\epsilon,\rho,1}[\Lambda_\epsilon(h,p,\zeta,\psi)]}{\partial \Lambda_\epsilon(h,p,\zeta,\psi)};$$

$$\frac{\partial \Lambda_\epsilon(h,p,\zeta,\psi)}{\partial h}=\frac{1}{2}\sum_{\iota=1}^{3}\sum_{\upsilon=1}^{\sigma}\left\{\partial\frac{\alpha_\iota^\upsilon(x(t_\upsilon|h,p),\zeta)}{\partial x(t_\upsilon|h,p)}\varpi(t_\upsilon|h,p)-\frac{\partial\beta_\iota^\upsilon(x(t_\upsilon|h,p),\zeta)}{\partial x(t_\upsilon|h,p)}\varpi(t_\upsilon|h,p)+\right.$$
$$\left.2\frac{\partial \gamma_{\epsilon,\rho,3\sigma}[\beta_\iota^\upsilon(x(t_\upsilon|h,p),\zeta)]}{\partial\beta_\iota^\upsilon(x(t_\upsilon|h,p),\zeta)}\frac{\partial\beta_\iota^\upsilon(x(t_\upsilon|h,p),\zeta)}{\partial x(t_\upsilon|h,p)}\varpi(t_\upsilon|h,p)\right\};$$

$$\frac{\partial \Lambda_\epsilon(h,p,\zeta,\psi)}{\partial p_\iota}=\frac{1}{2}\sum_{\iota=1}^{3}\sum_{\upsilon=1}^{\sigma}\left\{\partial\frac{\alpha_\iota^\upsilon(x(t_\upsilon|h,p),\zeta)}{\partial x(t_\upsilon|h,p)}\pi_\iota(t_\upsilon|h,p)-\frac{\partial\beta_\iota^\upsilon(x(t_\upsilon|h,p),\zeta)}{\partial x(t_\upsilon|h,p)}\pi_\iota(t_\upsilon|h,p)+\right.$$
$$\left.2\frac{\partial \gamma_{\epsilon,\rho,3\sigma}[\beta_\iota^\upsilon(x(t_\upsilon|h,p),\zeta)]}{\partial\beta_\iota^\upsilon(x(t_\upsilon|h,p),\zeta)}\frac{\partial\beta_\iota^\upsilon(x(t_\upsilon|h,p),\zeta)}{\partial x(t_\upsilon|h,p)}\pi_\iota(t_\upsilon|h,p)\right\},\iota\in I_9;$$

$$\frac{\partial \Lambda_\epsilon(h,p,\zeta,\psi)}{\partial \zeta_\iota^\upsilon}=\frac{1}{2}\left\{\frac{\partial\alpha_\iota^\upsilon(x(t_\upsilon|h,p),\zeta)}{\partial\zeta_\iota^\upsilon}-\frac{\partial\beta_\iota^\upsilon(x(t_\upsilon|h,p),\zeta)}{\partial\zeta_\iota^\upsilon}+2\frac{\partial \gamma_{\epsilon,\rho,3\sigma}[\beta_\iota^\upsilon(x(t_\upsilon|h,p),\zeta)]}{\partial\beta_\iota^\upsilon(x(t_\upsilon|h,p),\zeta)}\frac{\partial\beta_\iota^\upsilon(x(t_\upsilon|h,p),\zeta)}{\partial\zeta_\iota^\upsilon}\right\};$$

$$\frac{\partial\alpha_\iota^\upsilon(x(t_\upsilon|h,p),\zeta)}{\partial\zeta_\iota^\upsilon}=-(\frac{d_\iota^\upsilon}{\overline{d}_\iota^\upsilon}+1),\quad\frac{\partial\beta_\iota^\upsilon(x(t_\upsilon|h,p),\zeta)}{\partial\zeta_\iota^\upsilon}=-(\frac{d_\iota^\upsilon}{\overline{d}_\iota^\upsilon}-1).$$

Remark 2. *There exist two steps to calculate the gradients of the cost function $J_{\rho,\epsilon}^{\mathbf{H}}(h,p,\psi,\zeta)$ with respect to $\lambda\in\left\{h;p_\iota,\iota\in I_9;\psi;\zeta_\iota^\upsilon,\iota\in I_3,\upsilon\in I_\sigma\right\}$: (i) solve systems (2.1), (3.30), and (3.32) using a numerical differential scheme, progressing forward in time to obtain their respective solutions; (ii) compute the gradients based on Theorem 3. When these gradients are derived, a gradient-based NLP method can be applied to solve Problem **H**.*

3.6. Hybrid optimization algorithm

The numerical techniques for solving Problem **H** can be divided into two main categories: deterministic methods and intelligent methods [46]. Deterministic methods rely on known information,

such as gradients, to solve Problem **H**, and are recognized for their fast convergence. In contrast, intelligent methods often employ probabilistic strategies that are more effective for global optimization. While stochastic methods excel at exploring the global search space, they depend on random search processes, which can be time-consuming.

Typically, gradient-based methods [46] are applied to local optimization problems, as they are well-suited for refining solutions near the initial starting point. To address issues like getting stuck in local optima and premature convergence, a particle swarm optimization (PSO) algorithm, an example of an intelligent method, has been introduced [41].

To achieve a balance between local refinement and global exploration, a hybrid optimization algorithm, referred to as Algorithm 1, has been developed to effectively address Problem **H**. In this algorithm, the modified PSO method [41] is employed for the exploration phase, while gradient-based techniques are used for local exploitation, as noted in Line 3 of Algorithm 1. This hybrid approach combines the advantages of the PSO algorithm with the efficiency of gradient-based methods, which leverage the gradient information of the objective function in Problem **H**.

Let M_{sup} be a sufficiently large number. Let ρ_{sup} be a predefined parameter used to ensure the termination of the penalty approach. Let ρ^0 and ϵ^0 be the initial values of ρ and ϵ, respectively. Let ϵ be a predefined parameter used to ensure the termination of Algorithm 1. Let E_J be the convergence tolerance of the cost function. Let $\Psi_{\rho,\epsilon} := (h, p^{\mathrm{T}}, \psi, \zeta^{\mathrm{T}})^{\mathrm{T}}$. Let $\Psi^0_{\rho,\epsilon}$ be the initial value of $\Psi_{\rho,\epsilon}$.

Algorithm 1 Solving Problem **H**.

1: Choose initial values of $\Psi^0_{\rho,\epsilon}, \rho^0, \epsilon^0, M_{sup}, E_J, \rho_{sup}, \epsilon^{\min}$.
2: Set $J_{before} := M_{sup}, \rho := \rho^0, \epsilon := \epsilon^0$.
3: Solving Problem **H** with the cost function $J^{\mathbf{H}}_{\rho,\epsilon}(\Psi^0_{\rho,\epsilon})$ using the hybrid algorithm designed through the combination of PSO [41] and the gradient-based method [46] developed based on Theorem 3 to obtain the optimal solution, denoted by $\Psi^*_{\rho,\epsilon}$
4: Set $J_{present} := J^{\mathbf{H}}_{\rho,\epsilon}(\Psi^*_{\rho,\epsilon})$.
5: **if** $|J_{present} - J_{before}| < E_J$ **then**
6: **return** $\Psi^*_{\rho,\epsilon}$;
7: **else**
8: $J_{before} := J_{present}, \rho := 10\rho$.
9: **if** $\rho < \rho_{sup}$ **then**
10: Set $\Psi^0_{\rho,\epsilon} := \Psi^*_{\rho,\epsilon}$ and go to 3;
11: **else**
12: $J_{before} := J_{present}, \rho := \rho^0; \epsilon := 0.1\epsilon$.
13: **if** $\epsilon > \epsilon^{\min}$ **then**
14: Set $\Psi^0_{\rho,\epsilon} := \Psi^*_{\rho,\epsilon}$ and go to 3;
15: **else**
16: **return** $\Psi^*_{\rho,\epsilon}$.
17: **end if**
18: **end if**
19: **end if**

Remark 3. *Essentially, the goals of Algorithm 1 are for decreasing the value of ϵ to approximate the* max *function and increase the value of ρ to ensure that $|J_{present} - J_{before}| < E_J$ is satisfied.*

The convergence of Algorithm 1 is assured as stated in the following. (i) Theorem 1 guarantees that if the smoothing parameter ϵ is small enough, the solution of Problem **H** *is an approximate solution of Problem* **F**; *(ii) by virtue of the fundamental principles of the penalty function, when ρ is large enough, any solution to Problem* **F** *is a feasible solution to Problem* **E** *[40, 45]. Note that in Section 3.1, Problem* ***A*** *can be converted into Problem* **E**.

The selection of hyperparameters is often based on a combination of experience, experimentation, and evaluation metrics. For example, inertia weight (w): typically chosen between 0.4 and 0.9; acceleration factors (c_1, c_2): generally set between 1.5 and 2; population size: usually ranges from 20 to 50.

4. Numerical results

4.1. Experiment design

All computations are performed using MATLAB R2016a on a computer equipped with a 3.70 GHz Intel Core i9-10900K CPU and 32.0 GB of RAM. In the numerical computation, the cultivation conditions have been studied in [16]. During the fed-batch culture, alkali, which is intermittently put in the fermentor to maintain its pH value at approximately 7.0, can cause a chemical reaction with acetic acid. This operation significantly affects the correctness of the extracellular concentrations of acetic acid and ethanol. Therefore, in this paper, the focus is on the discrepancy between the measured output data and the system output for the first three substances. The fed-batch culture evolves starting from the initial states x_0.

The feeding process starts at $\tau^1 = 5.33$ (h). The feeding time starts at $\tau^{2i+1}, i \in \Lambda_2$ and stops at $\tau^{2i+2}, i \in \Lambda_1$. They were determined through the experiment reported in [43]. To obtain the value of the cost function (3.9) in Problem **F**, system (2.1) needs to be solved numerically. Thus, the computational load is rather large. Therefore, for saving the calculation time, the process of the fed-batch culture reported in [47] is grouped into ten stages (i.e., Batch Stage and Stages II–X). In each of Stages II–X, the time duration for each batch stage is 100s minus the duration of the feeding process. The feeding processes in Stages II–X last for 5, 7, 8, 7, 6, 4, 3, 2, 1, 2, and 1 seconds out of every 100 seconds, leaving 95, 93, 92, 93, 96, 97, 98, 99, 98, and 99 seconds for batch processes, respectively.

4.2. Parameters setting

Based on [43], the experiment parameters used in the calculation of the solution of system (2.1) are, respectively, $x_0 = (0.1115\ \text{gL}^{-1},\ 495\ \text{mmolL}^{-1},\ 0,\ 0,\ 0)^{\mathrm{T}}, \underline{x} = (0.0001, 0.1, 0, 0, 0)^{\mathrm{T}}, \overline{x} = (15, 2039, 939.5, 1026, 360.9)^{\mathrm{T}}$, $r = 0.75$, $C_{s0} = 10,762\ \text{mmolL}^{-1}$, $V_0 = 5$ L, $t_f = 26.8$ (h), $N = 783$, $\overline{h} = 12$ (hour), $\underline{p} = (0.438, 0.5, -7.35435, 0.0039, 33.845, 5.945445, 8.8648, 2.59, 10.225)^{\mathrm{T}}, \overline{p} = (1.314, 3.3, -2.45145, 0.0117, 101.535, 17.836335, 26.5944, 7.77, 90.675)^{\mathrm{T}}, \sigma = |\mathcal{T}| = 16, \mathcal{T} = \{0, 2, 4, 5.83, 7.83, 9.83, 11.83, 13.83, 15.83, 17.83, 19.83, 21.83, 23.83, 24.83, 25.83, 26.83\}$.

Euler method is used to solve system (2.1) with a step size of 1/3600. In Algorithm 1, the value of $\rho^0, \epsilon^0, M_{sup}, E_J, \rho_{sup}, \epsilon^{\min}$ are, respectively, $1, 10, 10^7, 10^{-2}, 10^6, 10^{-6}$. In addition, the value of $\Psi^0_{\rho,\epsilon}$ is selected in its feasible region. The initial function $\phi(t)$ is obtained through the use of the cubic spline method to interpolate the experimental data [40].

Remark 4. *Practitioners of numerical approximation are most concerned with* **truncation error**, *which in a numerical method is an error that is caused by using simple approximations to represent exact mathematical formulas. We seek information about errors on both a local and global scale.* **Local truncation error** *is the amount of truncation error that occurs in one step of a numerical approximation.* **Global truncation error** *is the amount of truncation error that occurs in the use of a numerical approximation to solve a problem.*

We obtain the global truncation error at step n in any numerical approximation of (2.1) by

$$E_{n+1} = [1 + \Delta t f_x^{\ell}(x(t), x(t-h), p)]E_n - \frac{(\Delta t)^2}{2}\ddot{x}(\tau), \text{ where } \tau \in [0, t_f], \Delta t = t_{n+1} - t_n.$$

The **local truncation error** *for Euler's method is* $-\frac{(\Delta t)^2}{2}\ddot{x}(\tau)$. *The local truncation error has two factors of* Δt, *and we say that it is* $O((\Delta t)^2)$. *Based on [48], we have*

$$|E_n| \leq \frac{M}{2K}(e^{Kt_n} - 1)\Delta t,$$

where

$$K = \max_{(t,x(t),x(t-h),p)\in[0,t_f]\times W\times W\times\mathcal{P}} |f_x^{\ell}(x(t), x(t-h), p)| < \infty,$$

$$M = \max_{(t,x(t),x(t-h),p)\in[0,t_f]\times W\times W\times\mathcal{P}} |(f_t^{\ell} + f^{\ell} f_x^{\ell})(x(t), x(t-h), p)| < \infty.$$

We can see from the above description that the error vanishes as $\Delta t \to 0$. *Therefore, it is appropriate to use the Euler method for solving differential equations with* $\Delta t = 1/3600$ (*hour*), *which is small enough for the requirement of microbial fermentation simulation.*

4.3. Results and analysis

By virtue of Algorithm 1, the obtained optimal values of time-delay and system parameters are, respectively, $h^* = 0.824237042558447$ and $p^* = (1.142666910517339, 0.500218711380598, -3.629735870518227, 0.005853192577210, 99.392558636711343, 6.427113527307207, 8.869263769688454, 2.659555099094625, 80.975777396140828)^{\mathrm{T}}$.

The optimal time delay in literature [43] was obtained using precise experimental data. In contrast, the optimal time delay presented in this paper is derived under the worst-case distribution of experimental data. Therefore, the optimal time delays obtained in the two cases differ.

The relative errors between the computed values $x_\iota(t_\upsilon|h^*, p^*), t_\upsilon \in \mathcal{T}, \iota \in I_3, \upsilon \in I_\sigma$, and the experimental data $y_\iota(t_\upsilon), t_\upsilon \in \mathcal{T}, \iota \in I_3, \upsilon \in I_\sigma$, obtained with and without (denoted by "C" and "B", respectively) and taking into consideration of the worst-case probability distribution of the experimental data (WCPDED) are listed in Table 2, where the relative errors are defined by

$$e_\iota := \frac{\sum_{\upsilon=1}^{\sigma} |x_\iota(t_\upsilon|h^*, p^*) - y_\iota(t_\upsilon)|}{\sum_{\upsilon=1}^{\sigma} |y_\iota(t_\upsilon)|}, \iota \in I_3.$$

Here, $x(t_\upsilon|h^*, p^*) := (x_1(t_\upsilon|h^*, p^*), \cdots, x_5(t_\upsilon|h^*, p^*))^{\mathrm{T}} \in \mathbb{R}^5_+$ denotes the solution of system (2.1) corresponding to the obtained optimal time-delay h^* and system parameters p^* with the initial concentration x_0, and $y(t_\upsilon) = (y_1(t_\upsilon), \cdots, y_3(t_\upsilon))^{\mathrm{T}} \in \mathbb{R}^3_+$ is the experiment data at the time point $t_\upsilon \in \mathcal{T}$.

Table 2 reveals that the relative error of 'C' exceeds that of 'B'. This discrepancy can be attributed to the fact that the relative error for 'C' is calculated based on the WCPDED. This WCPDED reflects greater variability and uncertainty of the distribution of the experimental data in the measurements for 'C', which naturally leads to a higher relative error. In contrast, surrounding 'B' without taking into consideration of the WCPDED contributes to its comparatively lower relative error. Understanding these distinctions is crucial for accurately interpreting the reliability of our findings.

Table 2. Relative errors between B and C.

Relative error	$e_1(\%)$	$e_2(\%)$	$e_3(\%)$
B	10.7042322473258	14.8532582883951	5.64026160583546
C	12.1565939669873	23.4868141782757	5.79013674616066

Under the obtained h^* and p^*, the curve of the concentration of the first three substances are drawn in Figures 2–4 (the horizontal axis stands for time; the vertical axis stands for the concentrations of biomass, glycerol, and 1,3-PD, respectively; A: experimental data; B: the concentration change curve of biomass, glycerol, and 1,3-PD without considering the WCPDED, respectively; C: the concentration change curve of biomass, glycerol, and 1,3-PD that takes into account the WCPDED, respectively). The analysis presented in Figures 2–4 illustrates a reasonable alignment between the computed values and the experimental data under the WCPDED framework. To facilitate comparison, we determined the relative errors between the computed concentrations of biomass, glycerol, and 1,3-PD against the experimental results, utilizing the optimal time-delay and system parameters outlined in [43]. It is important to note that, unlike the optimization problem discussed in [43], our approach to parameter estimation relies solely on the first-order moment of the experimental data, thereby simplifying the process. Moreover, Algorithm 1 is designed to operate effectively without requiring precise knowledge of the distribution of the experimental data. The entire calculation process took approximately 15,653 seconds.

Figures 2–4 demonstrate that the NCTSTD system, utilizing the parameters h^* and p^*, effectively captures the dynamics of the fed-batch culture process, even within the constraints of the WCPDED. This suggests that our model not only aligns well with observed behaviors but also provides a robust framework for understanding the intricacies of the cultivation process. The ability of the NCTSTD system to maintain accuracy under these conditions highlights its potential for applications in optimizing bioprocesses.

Note that our method is under the assumption that the first-order moment information of the experimental data can be accurately obtained. Obviously, it will be practically important to extend this method to cases where moment information is also uncertain. This will be an interesting further research direction.

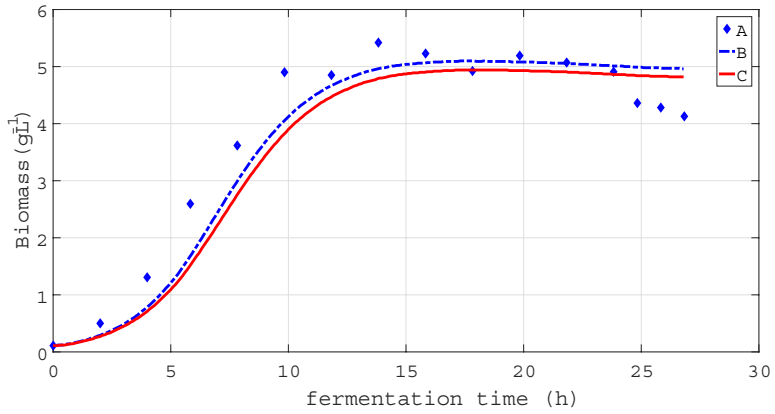

Figure 2. Comparison of experimental data and computed values of biomass concentrations under the value of h^* and p^*.

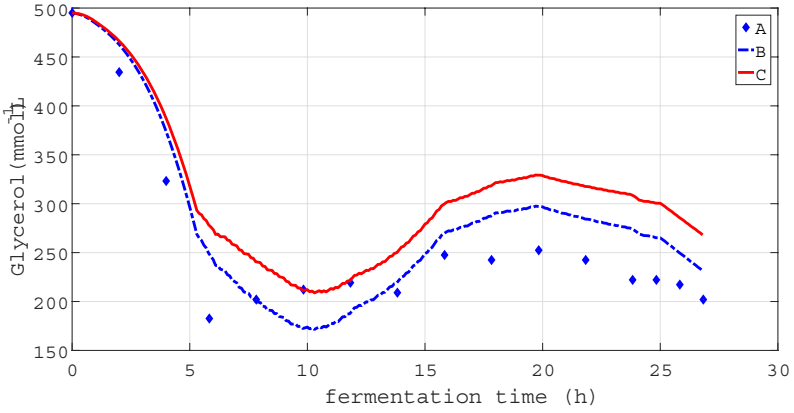

Figure 3. Comparison of experimental data and computed values of glycerol concentrations under the value of h^* and p^*.

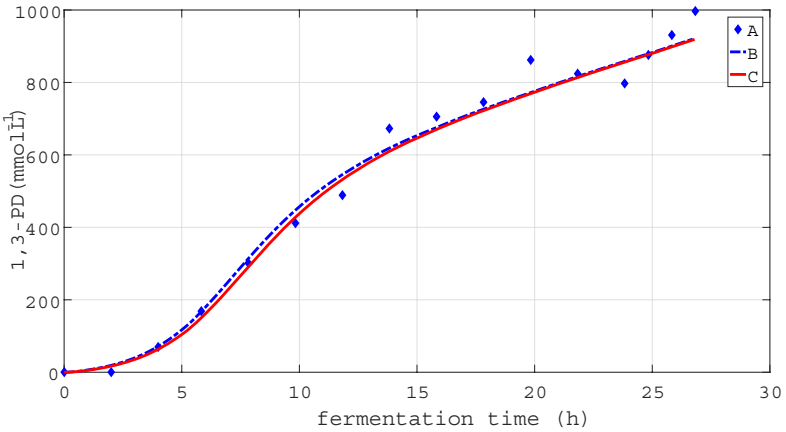

Figure 4. Comparison of experimental data and computed values of 1,3-PD concentrations under the the value of h^* and p^*.

5. Discussion

This paper studies an NCTSTD with unknown quantities, in which the measured output data is uncertain. The focus is to identify these unknown quantities. For this, we formulate the problem as a DRPE problem subject to the NCTSTD system. A hybrid algorithm is presented to solve the DRPE problem. Finally, numerical results are obtained, which demonstrate the validity of the designed algorithm.

On the other hand, it is important to note that the designed algorithm is based on the assumption that the first-order moment information of the measured output data with respect to its probability distribution is available, highlighting the need to extend the algorithm to cases where this moment information is uncertain. This would be an interesting further research area.

In summary, the designed algorithm provides a promising method to optimally identify time-delay and system parameters in the fed-batch culture process of glycerol bioconversion to 1,3-PD under the environment of uncertain measured output data. The potential extension of this algorithm to more general cases would enhance its scope of applications.

Use of AI tools declaration

The authors declare they have not used Artificial Intelligence (AI) tools in the creation of this article.

Acknowledgments

This work was supported in part by the National Key Research and Development Program of China under Grant 2022YFB3304600; in part by the Nature Science Foundation of Liaoning Province of China under Grant 2024-MS-015; in part by the Fundamental Research Funds for the Central Universities under Grant 3132024196, and Grant DUT22LAB305; in part by the National Natural Science Foundation of China under Grant 12271307, and Grant 12161076; in part by the China Postdoctoral Science Foundation under Grant 2019M661073; in part by the Xinghai Project of Dalian Maritime University.

Conflict of interest

The authors declare there is no conflicts of interest.

References

1. W. Liu, L. Yang, B. Yu, Kernel density estimation based distributionally robust mean-CVaR portfolio optimization, *J. Glob. Optim.*, **84** (2022), 1053–1077. https://doi.org/10.1007/s10898-022-01177-5
2. E. Delage, Y. Ye, Distributionally robust optimization under moment uncertainty with application to data-driven problems, *Oper. Res.*, **58** (2010), 595–612. https://doi.org/10.1287/opre.1090.0741
3. C. Peng, E. Delage, Data-driven optimization with distributionally robust second order stochastic dominance constraints, *Oper. Res.*, **72** (2024), 1298–1316. https://doi.org/10.1287/opre.2022.2387

4. S. Wang, L. Pang, H. Guo, H. Zhang, Distributionally robust optimization with multivariate second-order stochastic dominance constraints with applications in portfolio optimization, *Optimization*, **72** (2023), 1839–1862. https://doi.org/10.1080/02331934.2022.2048382

5. X. Tong, M. Li, H. Sun, Decision bounding problems for two-stage distributionally robust stochastic bilevel optimization, *J. Glob. Optim.*, **87** (2023), 679–707. https://doi.org/10.1007/s10898-022-01227-y

6. H. Xu, Y. Liu, H. Sun, Distributionally robust optimization with matrix moment constraints: Lagrange duality and cutting plane methods, *Math. Program.*, **169** (2018), 489–529. https://doi.org/10.1007/s10107-017-1143-6

7. B. Li, Y. Rong, J. Sun, K. L. Teo, A distributionally robust linear receiver design for multi-access space-time block coded MIMO systems, *IEEE Trans. Wireless Commun.*, **16** (2017), 464–474. https://doi.org/10.1109/TWC.2016.2625246

8. B. Li, Y. Tan, A. Wu, G. Duan, A distributionally robust optimization based method for stochastic model predictive control, *IEEE Trans. Autom. Control*, **67** (2022), 5762–5776. https://doi.org/10.1109/TAC.2021.3124750

9. A. Shapiro, Distributionally robust modeling of optimal control, *Oper. Res. Lett.*, **50** (2022), 561–567. https://doi.org/10.1016/j.orl.2022.08.002

10. B. P. Van Parys, D. Kuhn, P. J. Goulart, M. Morari, Distributionally robust control of constrained stochastic systems, *IEEE Trans. Autom. Control*, **61** (2015), 430–442. https://doi.org/10.1109/TAC.2015.2444134

11. A. Hakobyan, I. Yang, Wasserstein distributionally robust control of partially observable linear stochastic systems, *IEEE Trans. Autom. Control*, **69** (2024), 6121–6136. https://doi.org/10.1109/TAC.2024.3394348

12. B. Li, T. Guan, L. Dai, G. R. Duan. Distributionally robust model predictive control with output feedback. *IEEE Trans. Autom. Control*, **69** (2024), 3270–3277. https://doi.org/10.1109/TAC.2023.3321375

13. L. Huang, X. Zhou, N. Wu, Y. Sun, Z. Xiu, Selective extraction of 1,3-propanediol by phenylboronic acid-based ternary extraction system, *J. Chem. Technol. Biotechnol.*, **99** (2024), 1530–1540. https://doi.org/10.1002/jctb.7647

14. C. Groeger, W. Sabra, A. Zeng, Simultaneous production of 1,3-propanediol and n-butanol by Clostridium pasteurianum: In situ gas stripping and cellular metabolism, *Eng. Life Sci.*, **16** (2016), 664–674. https://doi.org/10.1002/elsc.201600058

15. J. Zhou, J. Shen, X. Wang, Y. Sun, Z. Xiu, Metabolism, morphology and transcriptome analysis of oscillatory behavior of Clostridium butyricum during long-term continuous fermentation for 1,3-propanediol production, *Biotechnol. Biofuels*, **13** (2020), 191. https://doi.org/10.1186/s13068-020-01831-8

16. S. Duan, Z. Zhang, X. Wang, Y. Sun, Y. Dong, L. Ren, et al., Co-production of 1,3-propanediol and phage phiKpS2 from the glycerol fermentation by Klebsiella pneumoniae, *Bioresour. Bioprocess.*, **11** (2024), 44. https://doi.org/10.1186/s40643-024-00760-w

17. L. Wang, J. Yuan, C. Wu, X. Wang, Practical algorithm for stochastic optimal control problem about microbial fermentation in batch culture, *Optim. Lett.*, **13** (2019), 527–541. https://doi.org/10.1007/s11590-017-1220-z

18. J. Wang, X. Zhang, J. Ye, J. Wang, E. Feng, Optimizing design for continuous conversion of glycerol to 1, 3-propanediol using discrete-valued optimal control, *J. Process Control*, **104** (2021), 126–134. https://doi.org/10.1016/j.jprocont.2021.06.010

19. J. Yuan, S. Lin, S. Zhang, C. Liu, Distributionally robust system identification for continuous fermentation nonlinear switched system under moment uncertainty of experimental data, *Appl. Math. Modell.*, **127** (2024), 679–695. https://doi.org/10.1016/j.apm.2023.12.023

20. J. Yuan, C. Wu, C. Liu, K. L. Teo, J. Xie, Robust suboptimal feedback control for a fed-batch nonlinear time-delayed switched system, *J. Franklin Inst.*, **360** (2023), 1835–1869. https://doi.org/10.1016/j.jfranklin.2022.12.027

21. C. Liu, G. Shi, G. Liu, D. Hu, Optimal control of a nonlinear state-dependent impulsive system in fed-batch process, *Int. J. Biomath.*, **16** (2023), 2350001. https://doi.org/10.1142/S1793524523500018

22. T. Niu, J. Zhai, H. Yin, E. Feng, C. Liu, The uncoupled microbial fed-batch fermentation optimization based on state-dependent switched system, *Int. J. Biomath.*, **14** (2021), 2150025. https://doi.org/10.1142/S179352452150025X

23. X. Li, M. Sun, Z. Gong, E. Feng, Multistage optimal control for microbial fed-batch fermentation process, *J. Ind. Manage. Optim.*, **18** (2022), 1709–1721. https://doi.org/10.3934/jimo.2021040

24. T. Niu, J. Zhai, H. Yin, E. Feng, Optimal control of nonlinear switched system in an uncoupled microbial fed-batch fermentation process, *J. Franklin Inst.*, **355** (2018), 6169–6190. https://doi.org/10.1016/j.jfranklin.2018.05.012

25. C. Zhang, S. Sharma, W. Wang, A. Zeng, A novel downstream process for highly pure 1,3-propanediol from an efficient fed-batch fermentation of raw glycerol by Clostridium pasteurianum, *Eng. Life Sci.*, **21** (2021), 351–363. https://doi.org/10.1002/elsc.202100012

26. J. Yuan, S. Zhao, D. Yang, C. Liu, C. Wu, T. Zhou, et al., Koopman modeling and optimal control for microbial fed-batch fermentation with switching operators, *Nonlinear Anal. Hybrid Syst.*, **52** (2024), 101461. https://doi.org/10.1016/j.nahs.2023.101461

27. J. Yuan, C. Wu, Z. Liu, S. Zhao, C. Yu, K. Teo, et al., Koopman modeling for optimal control of the perimeter of multi-region urban traffic networks, *Appl. Math. Modell.*, **138** (2025), 115742. https://doi.org/10.1016/j.apm.2024.115742

28. D. Wu, Y. Bai, C. Yu, A new computational approach for optimal control problems with multiple time-delay, *Automatica*, **101** (2019), 388–395. https://doi.org/10.1016/j.automatica.2018.12.036

29. M. Wang, N. Liu, Qualitative analysis and traveling wave solutions of a predator-prey model with time delay and stage structure, *Electron. Res. Arch.*, **32** (2024), 2665–2698. https://doi.org/10.3934/era.2024121

30. J. Yuan, C. Wu, K. Teo, S. Zhao, L. Meng, Perimeter control with state-dependent delays: optimal control model and computational method, *IEEE Trans. Intell. Transp. Syst.*, **23** (2022), 20614–20627. https://doi.org/10.1109/TITS.2022.3179729

31. D. Wu, Y. Chen, C. Yu, Y. Bai, K. Teo, Control parameterization approach to time-delay optimal control problems: A survey, *J. Ind. Manage. Optim.*, **19** (2023), 3750–3783. https://doi.org/10.3934/jimo.2022108

32. Y. Chen, X. Zhu, C. Yu, K. Teo, Sequential time scaling transformation technique for time-delay optimal control problem, *Commun. Nonlinear Sci. Numer. Simul.*, **133** (2024), 107988. https://doi.org/10.1016/j.cnsns.2024.107988

33. Z. Li, L. Pei, G. Duan, S. Chen, A non-autonomous time-delayed SIR model for COVID-19 epidemics prediction in China during the transmission of Omicron variant. *Electron. Res. Arch.*, **32** (2024), 2203–2228. https://doi.org/10.3934/era.2024100

34. J. Yuan, C. Wu, K. Teo, J. Xie, S. Wang, Computational method for feedback perimeter control of multiregion urban traffic networks with state-dependent delays, *Transp. Res. Part C Emerging Technol.*, **153** (2023), 104231. https://doi.org/10.1016/j.trc.2023.104231

35. J. Yuan, D. Yang, D. Xun, K. Teo, C. Wu, A. Li, et al., Sparse optimal control of cyber-physical systems via PQA approach, *Pac. J. Optim.*, Accepted, 2024.

36. C. Liu, R. Ryan, Q. Lin, K. Teo, Dynamic optimization for switched time-delay systems with state-dependent switching conditions, *SIAM J. Control Optim.*, **56** (2018), 3499–3523. https://doi.org/10.1137/16M1070530

37. C. Liu, C. Sun, Robust parameter identification of a nonlinear impulsive time-delay system in microbial fed-batch process, *Appl. Math. Modell.*, **111** (2022), 160–175. https://doi.org/10.1016/j.apm.2022.06.032

38. X. Gao, J. Zhai, E. Feng, Multi-objective optimization of a nonlinear switched time-delay system in microbial fed-batch process, *J. Franklin Inst.*, **357** (2020), 12609–12639. https://doi.org/10.1016/j.jfranklin.2020.07.036

39. C. Liu, M. Han, Time-delay optimal control of a fed-batch production involving multiple feeds, *Discrete Contin. Dyn. Syst. - S*, **13** (2020), 1697–1709. https://doi.org/10.3934/dcdss.2020099

40. J. Nocedal, S. J. Wright, *Numerical Optimization*, Springer, New York, 2006. https://doi.org/10.1007/978-0-387-40065-5

41. A. Banerjee, I. Abu-Mahfouz, A comparative analysis of particle swarm optimization and differential evolution algorithms for parameter estimation in nonlinear dynamic systems, *Chaos, Solitons Fractals*, **58** (2014), 65–83. https://doi.org/10.1016/j.chaos.2013.11.004

42. E. Anderson, P. Nash, *Linear Programming in Infinite-Dimensional Spaces: Theory and Applications*, Wiley, Chichester, United Kingdom, 1987. https://cir.nii.ac.jp/crid/1130000796325253632

43. C. Liu, Sensitivity analysis and parameter identification for a nonlinear time-delay system in microbial fed-batch process, *Appl. Math. Modell.*, **38** (2014), 1449–1463. https://doi.org/10.1016/j.apm.2013.07.039

44. P. Liu, X. Li, X. Liu, Y. Hu, An improved smoothing technique-based control vector parameterization method for optimal control problems with inequality path constraints, *Optim. Control. Appl. Methods*, **38** (2017), 586–600. https://doi.org/10.1002/oca.2273

45. X. Wu, J. Lin, K. Zhang, M. Cheng, A penalty function-based random search algorithm for optimal control of switched systems with stochastic constraints and its application in automobile test-driving with gear shifts, *Nonlinear Anal. Hybrid Syst.*, **45** (2022), 101218. https://doi.org/10.1016/j.nahs.2022.101218

46. K. Teo, B. Li, C. Yu, V. Rehbock, *Applied and Computational Optimal Control: A Control Parametrization Approach*, Springer, 2021. https://doi.org/10.1007/978-3-030-69913-0

47. X. Chen, D. J. Zhang, W. Qi, S. Gao, Z. Xiu, P. Xu, Microbial fed-batch production of 1, 3-propanediol by klebsiella pneumoniae under microaerobic conditions, *Appl. Microbiol. Biotechnol.*, **63** (2003), 143–146. https://doi.org/10.1007/s00253-003-1369-5

48. K. Atkinson, W. Han, D. Stewart, *Numerical Solution of Ordinary Differential Equations*, John Wiley and Sons, 2009. http://homepage.math.uiowa.edu/ atkinson/papers/NAODE_Book.pdf

AIMS Press