J. Nonlinear Var. Anal. 6 (2022), No. 2, pp. 35-64
Available online at http://jnva.biemdas.com
https://doi.org/10.23952/jnva.6.2022.2.04

MULTI-OBJECTIVE OPTIMIZATION OF A NONLINEAR BATCH TIME-DELAY SYSTEM WITH MINIMUM SYSTEM SENSITIVITY

LEI WANG[1], JINLONG YUAN[2,*], LIXIA MENG[1], SHUANG ZHAO[2], JUN XIE[3], MING HUANG[2], KOK LAY TEO[4,5]

[1]*Department of Mathematics, School of Science, Shihezi University, Shihezi 832003, China*
[2]*Department of Mathematics, School of Science, Dalian Maritime University, Dalian 116026, China*
[3]*Teaching and Research Office of Mathematics, Department of Basics, PLA Dalian Naval Academy, Dalian 11601, China*
[4]*School of Mathematical Sciences, Sunway University, Bandar Sunway 47500, Malaysia*
[5]*The Coordinated Innovation Center for Computable Modeling in Management Science, Tianjin University of Finance and Economics, Tianjin 300222, China*

Abstract. In this paper, we consider a nonlinear time-delay dynamic (NTDD) system with uncertain time-delay in batch culture of glycerol bioconversion to 1,3-propanediol (1,3-PD) induced by Klebsiella pneumoniae. Our goal is to design an optimization scheme for the NTDD system with the aim of balancing two competing objectives: (i) system cost (the relative error between experimental data and the output of the mathematical model); (ii) system sensitivity (the variation of the system cost with respect to uncertain time-delay). Thus, a multi-objective optimization problem (MOOP) governed by the NTDD system and subject to continuous state inequality constraints is proposed, where the two competing objective functions are to be minimized. The optimization variables in this problem are the initial concentrations of biomass and glycerol along with the free terminal time. The MOOP is first converted into a sequence of single-objective optimization problems (SOOCPs) by using convex weighted sum and modified normal boundary intersection methods. By incorporating the time scaling transformation, the constraint transcription and locally smoothing approximation, a parallel hybrid SOOCP solver is developed based on gradient-based method and genetic algorithm. Finally, numerical results are provided to verify the effectiveness of the proposed solution method.

Keywords. Multi-objective optimization; Nonlinear time-delay system; Parallel hybrid optimization; System sensitivity.

1. INTRODUCTION

1,3-propanediol (1,3-PD), a bifunctional chemical reagent, is used as a monomer for polyester and polyurethanes [1,2]. Compared with chemical synthesis, the bioconversion of glycerol induced by Klebsiella pneumoniae (*K. pneumoniae*) to 1,3-PD is much more attractive. This is because it is friendly to environment and can make use of cheap renewable feedstock [3]. Bioprocesses involved in producing different pharmaceutical products may be divided into three categories: batch culture (e.g. [4]), continuous culture (e.g. [5, 6]), and fed-batch culture (e.g. [7–9]). The batch culture (see [7]) is particularly

*Corresponding author.
E-mail address: yuanjinlong0613@163.com (J. Yuan).
Received October 9, 2021; Accepted December 20, 2021.

important due to the following reasons: (i) Since the batch culture can be expressed as an excessive metabolism process of glycerol, 1,3-PD yield, which is the ratio between the formation of 1,3-PD and the consumption of glycerol, is high. (ii) Batch culture is a simple operation mode when compared with fed-batch culture or continuous cultures. (iii) Batch culture is a basic operation during the process of controlling fed-batch and continuous cultures. For these reasons, batch culture has been actively studied in the existing literature, such as sensitivity analysis [10], stability [11], pathway identification [12], and joint estimation [13]. It has also been modeled as stochastic system [14], hybrid system [15], and nonlinear time-varying process [16], just to name a few.

The effect of time-delays has rarely been explored, however, its importance should not be ignored [17–20]: (i) a cell obliges to go through some growth process for which it requires some time before the reaction takes place; (ii) it takes certain time for the substrate together with the products to be transported across the cell membrane from outside to inside; and (iii) occasionally, due to the lack of knowledge or for the purpose of reducing complexity, a multitude of intermediate steps is omitted in the response system and hence a time delay is induced. Consequently, the optimal control of a batch fermentation process with time-delay and free terminal time and cost sensitivity constraint was investigated in [21] and the parameter identification for a nonlinear enzyme-catalytic dynamic system with time-delays was studied in [22]. However, the uncertain nature of time-delay was not taken into consideration. Clearly, the time-delay varied during the life span of the system [23–25]. Therefore, system sensitivity, which measures the change of the system cost in response to small variation of time-delay, should be considered.

There is always a trade-off between the system cost and the system sensitivity. In the existing literature, only a single objective, which is the system cost, is minimized in batch culture; see, for example, [10–15,22,26,27]. However, the optimal system cost may be sensitive to the variation of some of the time-delays and these time-delays are likely to change during the life span of the system. Thus, the trade-off between system cost and system sensitivity should be investigated and this becomes a multi-objective optimization problem (MOOP). In passing, it is noted that multiple objectives have been considered in optimizing biochemical processes, see, for example, [28–33]. For MOOPs, its aim is to find the solution set, called Pareto set. In Pareto set, there exists no solution which will improve one objective function without worsening at least one of the other objective functions [34]. There are roughly two types of methods for generating the Pareto set: (i) vector based approaches; and (ii) scalarization approaches [35]. For the vector based approach, the MOOPs are dealt with directly through the exploitation of population based stochastic methods. These methods are capable of producing the Pareto set directly. For the scalarization approach, a MOOP is converted into a single-objective optimization problem (SOOP) through the exploitation of additional scalar variables. For this approach, the Convex Weighted Sum (CWS) is the most common used method [36]. However, the CWS suffers from some intrinsic shortcomings, and consequently, some advanced scalarization approaches were introduced, such as: (i) the (Modified) Normal Boundary Intersection (mNBI); and (ii) the (Enhanced) Normalized Normal Constraint (eNNC) [37]. For the first approach, gradient-based algorithms can be used to find (at least locally) optimal solutions for large-scale and highly constrained MOOPs efficiently. However, for all the papers mentioned above, they do not consider the coexistence of factors, such as time-delay uncertainty and system sensitivity in the batch culture.

In this paper, a nonlinear time-delay dynamic (NTDD) system is modelled to describe batch bio-bioconversion process of glycerol to 1,3-PD induced by *K. pneumoniae*. Both the system cost (the relative error between experimental data and the output of the system model) and the system sensitivity (the variation of the system cost with respect to uncertain time-delay) are regarded as the objective functions. These two objective functions are conflicting. Therefore, a MOOP is formulated where the initial concentrations of biomass and glycerol together with the free terminal time are taken as the control variables. The MOOP is transcribed into a sequence of SOOPs by using the CWS and the mNBI methods. However, since the resulting SOOPs contains free terminal time, the existing single-objective

solvers [30–32] may fail to solve these SOOPs directly. Furthermore, to ensure that the concentrations of biomass, glycerol, 1,3-PD, acetate and ethanol are to lie within specified limits, appropriate continuous state inequality constraints are required to be imposed to the MOOP governed by the NTDD system. For these reasons, a new single-objective solver is developed to solve these resulting SOOPs. First, each of these SOOPs is approximated as a sequence of nonlinear mathematical programming sub-problems through the application of the time-scaling transformation [26], the constraint transcription, and the locally smoothing approximation techniques [26]. Then, the gradients of the objective functions and the constraint functions with respect to these decision variables are derived. On this basis, the new gradient-based single-objective solver is developed to solve each of these SOOPs. Numerical results are provided to demonstrate the effectiveness of the proposed solution method.

The rest of this paper is organized as follows. In Section 2, a nonlinear time-delay dynamic system is formulated to describe the process of batch culture. In Section 3, a MOOP is presented. In Section 4, a numerical solution method is developed. In Section 5, a numerical example is solved and the results obtained are discussed. In the last section, Section 6, we draw some conclusions.

2. Nonlinear Time-delay Dynamic System

The dynamic system of the batch fermentation process is formulated based on the following assumptions.

- H1. No medium is pumped into and extracted from the bioreactor during the process; and
- H2. The concentrations within bioreactor are uniformly distributed.

Under Hypothesis **H1** and **H2**, the mass balance relationships for biomass, substrate, and products in batch fermentation can be described [38] by

$$\begin{cases} \dot{x}(t) = f(x(t), x(t-\tau)) = \begin{Bmatrix} \mu(t)x_1(t-\tau), \\ -q_2(t)x_1(t-\tau), \\ q_3(t)x_1(t-\tau), \\ q_4(t)x_1(t-\tau), \\ q_5(t)x_1(t-\tau), \end{Bmatrix}, t \in (0, T], \\ x(t) = \phi(t),\ t \le 0, \\ x(0) = x_0, \end{cases} \tag{2.1}$$

where τ is an uncertain delay argument; $x(t) := [x_1(t), \ldots, x_5(t)]^\intercal \in \mathbb{R}^5$ is the continuous state vector (whose components are the state variables) with $x_1(t), \ldots, x_5(t)$ being, respectively, the concentrations of biomass, glycerol, 1,3-PD, acetate, and ethanol at time $t \in [0, T]$. In real world situation, the time delay argument is uncertain and cannot be determined exactly. Some approaches were proposed in [38] for the estimation of the value of the uncertain time delay argument. We will use the estimate (i.e., $\tau = 0.26$) as the nominal value of the time delay argument in this paper.

In system (2.1), the specific growth rate of cells $\mu(t)$ is defined by

$$\mu(t) = \frac{\mu_m x_2(t)}{x_2(t) + k_1} \prod_{j=2}^{5} \left(1 - \frac{x_j}{x_j^*}\right). \tag{2.2}$$

In system (2.1), the specific consumption rate of substrate $q_2(t)$ is defined by

$$q_2(t) = m_2 + \frac{\mu(t)}{Y_2}. \tag{2.3}$$

In system (2.1), the specific formation rates of 1,3-PD $q_3(t)$ is defined by

$$q_3(t) = -m_3 + \mu(t) Y_3. \tag{2.4}$$

TABLE 1. Definitions of variables in Section 2.

Variable	Representation	Variable	Representation
I_n	the set $\{1,2,\ldots,n\}$	$\mathbb{R}_+$	the set of nonnegative real numbers
$\mathbb{R}$	the set of real numbers	$A^\intercal$	the transposition of matrix or vector A
a_1, b_1	the lower and upper bounds of the initial state x_{01}	a_2, b_2	the lower and upper bounds of initial state x_{02}
a_3, b_3	the lower and upper bounds of time durations	$T \in [a_3, b_3]$	the free terminal time
$\mathscr{U}$	$[a_1, b_1] \times [a_2, b_2]$	χ	$[x_{01}, x_{02}]^\intercal \in \mathbb{R}^2_+$
$x_0 := [x_{01}, \ldots, x_{05}]^\intercal$	the initial state vector	$\phi(t)$	a given initial function
μ_m	the maximum specific growth rate	k_1	the Monod saturation constant
m_2	the maintenance term of substrate consumption	Y_2	the maximum growth yields
m_3, m_4, m_5	the maintenance terms of 1,3-PD, acetate and ethanol	Y_3, Y_4, Y_5	the maximum 1,3-PD, acetate and ethanol yields

In system (2.1), the specific formation rates of acetic acid $q_4(t)$ is defined by

$$q_4(t) = -m_4 + \mu(t)Y_4. \tag{2.5}$$

In system (2.1), the specific formation rates of ethanol $q_5(t)$ is defined by

$$q_5(t) = m_5 + \mu(t)Y_5. \tag{2.6}$$

The definition of the kinetic parameters appeared in formulas (2.2)-(2.6) are listed in Table 2.

For system (2.1), there exists a unique continuous solution $x(\cdot|\chi,T,\tau) \in \mathbb{R}^5$ corresponding to the pair of decision variables (χ,T) and the nominal time-delay τ on $(-\infty,+\infty)$. For the five types of substance, there exist critical concentrations beyond which the cells will stop growing. Thus, we will impose restrictions on the concentrations of these five types of substance. The lower bound concentrations x_* of $x(t)$ for cells growth are given in [38], i.e.,

$$x_* := [x_{1*},\cdots,x_{5*}]^\intercal = [0.01,0,0,0,0]^\intercal, \tag{2.7}$$

and the upper bound concentrations x^* of $x(t)$ for cells growth are given in [38], i.e.,

$$x^* := [x_1^*,\cdots,x_5^*]^\intercal = [15,2039,939.5,1026,360.9]^\intercal. \tag{2.8}$$

In view of the results obtained in [38], the feasible set W of the state vector $x(t)$ is defined by

$$x(t|\chi,T,\tau) \in W := [x_*,x^*] = \prod_{i=1}^{5}[x_{i*},x_i^*] \subset \mathbb{R}_+^5, t \in [0,T].$$

3. Multi-objective Optimization Problem

In this section, a MOOCP governed by system (2.1) with uncertain time delay is formulated, where the productivity of 1,3-PD is maximized, while the consumption rate of glycerol is minimized. Clearly, these two objectives are conflicting. They are, respectively, given by

(i) Maximizing the productivity of 1,3-PD:

$$\max G_1(\chi,T|\tau),$$

where

$$G_1(\chi,T|\tau) = \frac{x_3(T|\chi,T,\tau)}{T}.$$

(ii) Minimizing the consumption rate of glycerol:

$$\min G_2(\chi,T|\tau),$$

where

$$G_2(\chi,T|\tau) = \frac{u_2 - x_2(T|\chi,T,\tau)}{T}.$$

Let

$$J(\chi,T|\tau) = (-G_1(\chi,T|\tau), G_2(\chi,T|\tau))^\intercal$$

be the system cost to be minimized.

For a real fermentation process, the following *continuous state inequality constraints* [39] are required to be satisfied at each $t \in [0,T]$

$$\begin{aligned} g_j(x(t|\chi,T,\tau)) &:= x_j(t|\chi,T,\tau) - x_j^* \le 0, \forall t \in [0,T], j \in I_5, \\ g_{j+5}(x(t|\chi,T,\tau)) &:= x_{j*} - x_j(t|\chi,T,\tau) \le 0, \forall t \in [0,T], j \in I_5, \end{aligned} \tag{3.1}$$

TABLE 2. The value of the kinetic parameters appeared in formulas (2.2)-(2.6), which were listed in [38].

Symbol	Value	Unit	Symbol	Value	Unit	Symbol	Value	Unit	Symbol	Value	Unit
μ_m	0.994	h^{-1}	k_1	0.368	$mmolL^{-1}$	m_2	3.24	$mmolg^{-1}h^{-1}$	m_3	3.679	$mmolg^{-1}h^{-1}$
m_4	0.491	$mmolg^{-1}h^{-1}$	m_5	7.309	$mmolg^{-1}h^{-1}$	Y_2	0.085	$mmolg^{-1}$	Y_3	76	$mmolg^{-1}$
Y_4	35.54	$mmolg^{-1}$	Y_5	14.78	$mmolg^{-1}$						

where $x_j^*, j \in I_5$ and $x_{j*}, j \in I_5$, are defined by (2.7)-(2.8). Then, by incorporating constraint (3.1), we can formulate the following problem:

$$\begin{aligned} \textbf{MOOP}: \quad \min \quad & J(\chi, T|\tau) \\ s.t. \quad & g_j(x(t|\chi,T,\tau)) \le 0,\ \forall t \in [0,T], j \in I_{10}, \\ & (\chi,T) \in \mathscr{U} \times [a_3,b_3]. \end{aligned}$$

However, for **MOOP**, the estimation of the time-delay τ is inaccurate. We then consider the sensitivity of the objective vector with respect to the variation of τ under the optimal (χ^*, T^*) obtained. To address this issue, inspired by the work in [23–25, 40, 41], the following measure of the productivity sensitivity with respect to the variation of τ is considered:

$$G_3(\chi,T|\tau) = \left(\frac{\partial G_1(\chi,T|\tau)}{\partial \tau}\right)^2. \tag{3.2}$$

Obviously, (3.2) measures the changes of the productivity with reference to small variation of τ. Therefore, a low value of the productivity sensitivity implies that the productivity of 1,3-PD is insensitive with reference to the variation of τ. The measure of the sensitivity of the consumption rate with respect to the variation of τ is given by

$$G_4(\chi,T|\tau) = \left(\frac{\partial G_2(\chi,T|\tau)}{\partial \tau}\right)^2. \tag{3.3}$$

Clearly, (3.3) measures the changes of the consumption rate with reference to small variation of τ. Likewise, a low value of the sensitivity of the consumption rate indicates that the consumption rate of glycerol is insensitivity with reference to the variation of τ. We may now define the following modified objective vector $J^\beta(\chi,T|\tau) = (-G_1(\chi,T|\tau)+\beta G_3(\chi,T|\tau), G_2(\chi,T|\tau)+\beta G_4(\chi,T|\tau))^\intercal$, where $\beta > 0$ is a weight coefficient. The weight β controls the relative importance between $-G_1(\chi,T|\tau)$ and $G_3(\chi,T|\tau)$ as well as between $G_2(\chi,T|\tau)$ and $G_4(\chi,T|\tau)$, respectively. If β is equal to zero, then the priority is to minimize $-G_1(\chi,T|\tau)$ and $G_2(\chi,T|\tau)$. When the distribution of the time delay τ is known exactly, $\beta = 0$ is the best option for minimizing $-G_1(\chi,T|\tau)$ and $G_2(\chi,T|\tau)$. However, when there are errors and/or uncertainties in the distribution of the time delay τ, it is essential to choose $\tau > 0$ to ensure solution robustness.

Hence, **MOOP** can be formulated as:

$$\begin{aligned} \textbf{MOOP}^\beta: \quad \min \quad & J^\beta(\chi, T|\tau) \\ s.t. \quad & g_j(x(t|\chi,T,\tau)) \le 0,\ \forall t \in [0,T], j \in I_{10}, \\ & (\chi,T) \in \mathscr{U} \times [a_3,b_3]. \end{aligned}$$

4. Computational Approaches

Note that **MOOP**$^\beta$ becomes **MOOP** when the weight coefficient β is zero. Note that **MOOP**$^\beta$ has four non-standard features: (i) the terminal time is free rather than fixed; (ii) the objective function contains sensitivity terms; (iii) the objective function is not a scalar but a vector; and (iv) constraints (3.1) are continuous state inequality constraints required to be satisfied at each $t \in [0,T]$.

4.1. Hybrid time-scaling transformation.

For **MOOP**$^\beta$, there are two types of direct approaches to compute the sensitivity measures defined by (3.2) and (3.3): (i) the variational method; and (ii) the costate method. However, it remains a challenge to deal with the free terminal time T [42]. To circumvent this challenge, the time-scaling transformation, which transforms **MOOP**$^\beta$ into an equivalent form with fixed terminal time, was proposed in [43, 44]. However, the transformation will cause the fixed time delay in the original time interval to become a variable time-delay in the new time interval. Accordingly, this

time-scaling transformation complicates, rather than simplifies, the computation of **MOOP**$^{\beta}$ subject to the system (2.1) with time-delay. In [45], a novel hybrid time-scaling transformation was presented. Details are given below.

This transformation takes effect through introducing a new time variable s and relating s to t through the equation

$$t = \bar{\mu}(s|T) = \begin{cases} T, & \text{if } s = 1, \\ T \cdot s, & \text{if } 0 \leq s < 1, \\ s, & \text{if } s < 0, \end{cases}$$

where $\bar{\mu}$ is referred to as a *time-scaling function*.

Clearly, $\bar{\mu}$ is non-decreasing, continuous, and piecewise-linear. Furthermore,

$$\frac{\mathrm{d}\bar{\mu}(s|T)}{\mathrm{d}s} = T, s \in (0,1).$$

Let $\tilde{x}(s) := x(\bar{\mu}(s|T))$. Then, we obtain the following new nonlinear system through the application of the time-scaling transformation $t = \bar{\mu}(s|T)$:

$$\begin{cases} \dot{\tilde{x}}(s) = \dfrac{\mathrm{d}}{\mathrm{d}s}\big(\tilde{x}(s)\big) = \dfrac{\mathrm{d}}{\mathrm{d}s}\big(x(\bar{\mu}(s|T))\big) = T\dfrac{\mathrm{d}x(\bar{\mu}(s|T))}{\mathrm{d}t}, s \in (0,1], \\ \tilde{x}(s) = \tilde{\phi}(s) = \phi(\bar{\mu}(s|T)), s \leq 0, \\ \tilde{x}(0) = x_0, s = 0. \end{cases} \tag{4.1}$$

Note that the time delayed argument $\bar{\mu}(s|T) - \tau$ is related to both s and T. Let $s_{delay} = s_{delay}(s,T,\tau)$ and $t_{delay} = t_{delay}(s,T,\tau)$ denote the delayed time points in the new time interval $[0,1]$ and the original time interval $[0,T]$, respectively. The relationship between s_{delay} and t_{delay} is given by

$$\bar{\mu}(s_{delay}|T) = t_{delay} = \bar{\mu}(s|T) - \tau, s \in [0,1]. \tag{4.2}$$

Hence, for $s \geq 0$, we obtain

$$x(t-\tau) = x(\bar{\mu}(s|T) - \tau) = x(\bar{\mu}(s_{delay}|T)) = \tilde{x}(s_{delay}). \tag{4.3}$$

Substituting (2.1) into (4.3) and (4.1), we see that

$$\begin{cases} \dot{\tilde{x}}(s) = \tilde{f}(\tilde{x}(s), \tilde{x}(s_{delay}), T) = Tf(\tilde{x}(s), \tilde{x}(s_{delay})), \ \bar{\mu}(s_{delay}|T) \in (0,T], s \in (0,1], \\ \tilde{x}(s) = \tilde{\phi}(s) = \phi(\bar{\mu}(s|T)), s \leq 0, \\ \tilde{x}(0) = x_0, s = 0. \end{cases} \tag{4.4}$$

The difficulty of solving (4.4) is that s_{delay} relies on s and T.

- **Case 1:** When $\bar{\mu}(s_{delay}|T) < 0$, it follows from (4.1) that $s_{delay} < 0$, thus yielding

$$s_{delay} = \bar{\mu}(s_{delay}|T) = \bar{\mu}(s|T) - \tau. \tag{4.5}$$

- **Case 2:** When $\bar{\mu}(s_{delay}|T) \geq 0$, it follows from (4.1) and (4.2) that s_{delay} satisfies:

$$T \cdot s - \tau = T \cdot s_{delay}. \tag{4.6}$$

To obtain the value of the delayed state $\tilde{x}(s_{delay})$, the delay time s_{delay} is required to be obtained. It should be noted that s_{delay} is a variable, and its value is obtained through solving (4.6) at each time point s. Thus, to calculate only a single objective function value, we need to solve this equation at least T/h times, where h (usually taken as $1/3600$) is the maximum step length, when solving the time-delay differential equations. Obviously, the numerical solution of system (4.4) defined on the new time horizon is computationally very demanding. Consequently, we are required to find a better way to get the value of $\tilde{x}(s_{delay})$ and then solve (4.6).

To overcome the computational burden of calculating s_{delay}, the following identity is applied:

$$\tilde{x}(s_{delay}) = x(\bar{\mu}(s_{delay}|T)) = x(t_{delay}). \tag{4.7}$$

Substituting (4.2) and (4.7) into (4.4) yields the following dynamic system:

$$\begin{cases} \dot{\tilde{x}}(s) = \tilde{f}(\tilde{x}(s), x(t_{delay}), T) = T f(\tilde{x}(s), x(t_{delay})),\ t_{delay} \in (0,T], s \in (0,1], \\ \tilde{x}(s) = \tilde{\phi}(s) = \phi(\bar{\mu}(s|T)), s \le 0, \\ \tilde{x}(0) = x_0, s = 0. \end{cases} \tag{4.8}$$

Let $\tilde{x}(\cdot|\chi,T,\tau)$ denote the solution of system (4.8) corresponding to the admissible pair (χ,T).

There is no difficulty in evaluating the delay t_{delay} (unlike s_{delay}) in system (4.8), as it can be obtained explicitly by using $t_{delay} := \bar{\mu}(s|T) - \tau$. However, the right-hand side of system (4.8) depends on two coupled time-delay systems. One defined on the original time scale and the other defined on the new time scale. Thus, this new transformation, named *hybrid time-scaling transformation*, was proposed in [45]. Its construction is given below.

For a given pair $(\chi,T) \in \mathscr{U} \times [a_3,b_3]$, $\tilde{x}(\cdot|\chi,T,\tau)$ can be obtained as follows:

- **Case 1:** For $s \le 0$, based on the initial condition of system (4.8), we have $\tilde{x}(s|\chi,T,\tau) = \phi(s)$.
- **Case 2:** For $s > 0$, $\tilde{x}(s|\chi,T,\tau)$ is determined by solving system (4.8) numerically. When solving this differential equation, $x(t_{delay})$ defined by the right-hand side of system (4.8) is obtained using the former values of $\tilde{x}(s|\chi,T,\tau)$. More explicitly, to obtain the state value at $t_{delay} := \bar{\mu}(s|T) - \tau$, we first find the unique past time point t_ℓ on the original time interval such that $t_\ell \le t_{delay} \le t_{\ell+1}$. These knot points must exist due to the fact that $\bar{\mu}(s|T)$ is continuous and strictly increasing on $(-\infty, 1]$. The value of $x(t_{delay})$ can be determined by interpolating the values of $x(t_\ell)$ and $x(t_{\ell+1})$.

Correspondingly, $G_j(\chi,T|\tau), j \in I_4, J^\beta(\chi,T|\tau)$ and formula (3.1) together with **MOOP**$^\beta$ are, respectively, transformed into:

$$\widetilde{G}_1(\chi,T|\tau) = \frac{\tilde{x}_3(1|\chi,T,\tau)}{T},\ \widetilde{G}_2(\chi,T|\tau) = \frac{u_2 - \tilde{x}_2(1|\chi,T,\tau)}{T},$$

$$\widetilde{G}_3(\chi,T|\tau) = \left(\frac{\partial \widetilde{G}_1(\chi,T|\tau)}{\partial \tau}\right)^2,\ \widetilde{G}_4(\chi,T|\tau) = \left(\frac{\partial \widetilde{G}_2(\chi,T|\tau)}{\partial \tau}\right)^2,$$

$$\begin{aligned} \widetilde{J}^\beta(\chi,T|\tau) &= (\widetilde{J}_1^\beta(\chi,T|\tau), \widetilde{J}_2^\beta(\chi,T|\tau))^\intercal \\ &:= (-\widetilde{G}_1(\chi,T|\tau) + \beta \widetilde{G}_3(\chi,T|\tau), \widetilde{G}_2(\chi,T|\tau) + \beta \widetilde{G}_4(\chi,T|\tau))^\intercal, \end{aligned} \tag{4.9}$$

$$\widetilde{g}_j(\tilde{x}(s|\chi,T,\tau)) := \tilde{x}_j(s|\chi,T,\tau) - x_j^* \le 0, \forall s \in [0,1], j \in I_5, \tag{4.10}$$

$$\widetilde{g}_{j+5}(\tilde{x}(s|\chi,T,\tau)) := x_{j*} - \tilde{x}_j(s|\chi,T,\tau) \le 0, \forall s \in [0,1], j \in I_5, \tag{4.11}$$

and

$$\widetilde{\textbf{MOOP}}^\beta : \begin{array}{ll} \min & \widetilde{J}^\beta(\chi,T|\tau) \\ s.t. & \widetilde{g}_j(\tilde{x}(s|\chi,T,\tau)) \le 0,\ \forall s \in [0,1], j \in I_{10}, \\ & (\chi,T) \in \mathscr{U} \times [a_3,b_3]. \end{array}$$

4.2. **Sensitivity computation.** We now derive the formulas for the productivity sensitivity and the sensitivity of the consumption rate of $\widetilde{\textbf{MOOP}}^\beta$ in the following theorem.

Theorem 4.1.

$$\widetilde{G}_3(\chi,T|\tau) = \left(\frac{\tilde{\psi}_3(1|\chi,T,\tau)}{T}\right)^2,\ \widetilde{G}_4(\chi,T|\tau) = \left(\frac{\tilde{\psi}_2(1|\chi,T,\tau)}{T}\right)^2,$$

where $\tilde{\psi}(s|\chi,T,\tau)=\big(\tilde{\psi}_1(s|\chi,T,\tau),\ldots,\tilde{\psi}_5(s|\chi,T,\tau)\big)^{\mathsf{T}}\in\mathbb{R}^5$ *is the solution of the following auxiliary delay-differential system:*

$$\begin{aligned}\dot{\tilde{\psi}}(s|\chi,T,\tau)=&\frac{\partial\tilde{f}(\tilde{x}(s),x(t_{delay}),T)}{\partial\tilde{x}}\tilde{\psi}(s|\chi,T,\tau)\\&+\frac{\partial\tilde{f}(\tilde{x}(s),x(t_{delay}),T)}{\partial x(t_{delay})}\Big[\tilde{\psi}(s_{delay}|\chi,T,\tau)+\frac{\partial x(t_{delay})}{\partial s_{delay}}\frac{\partial s_{delay}}{\partial\tau}\Big],s\in(0,1],\end{aligned}\tag{4.12}$$

with

$$\tilde{\psi}(s|\chi,T,\tau)=\mathbf{0}\in\mathbb{R}^5,\quad\forall s\le 0.\tag{4.13}$$

Proof. Let $(\chi,T)\in\mathscr{U}\times[a_3,b_3]$ be an admissible pair. By [45], we have

$$\frac{\partial\tilde{x}_j(1|\chi,T,\tau)}{\partial\tau}=\tilde{\psi}_j(1|\chi,T,\tau),j\in I_5,$$

where $\tilde{\psi}(s|\chi,T,\tau)=\big(\tilde{\psi}_1(s|\chi,T,\tau),\ldots,\tilde{\psi}_5(s|\chi,T,\tau)\big)^{\mathsf{T}}\in\mathbb{R}^5$ is the solution of the systems (4.12) with initial condition (4.13). Thus, differentiating $\widetilde{G}_1(\chi,T|\tau)$ and $\widetilde{G}_2(\chi,T|\tau)$ with respect to τ yields

$$\frac{\partial\widetilde{G}_1(\chi,T|\tau)}{\partial\tau}=\frac{\partial}{\partial\tau}\Big\{\frac{\tilde{x}_3(1|\chi,T,\tau)}{T}\Big\}=\frac{1}{T}\frac{\partial\tilde{x}_3(1|\chi,T,\tau)}{\partial\tau}=\frac{\tilde{\psi}_3(1|\chi,T,\tau)}{T},$$

and

$$\frac{\partial\widetilde{G}_2(\chi,T|\tau)}{\partial\tau}=\frac{\partial}{\partial\tau}\Big\{-\frac{\tilde{x}_2(1|\chi,T,\tau)}{T}\Big\}=-\frac{1}{T}\frac{\partial\tilde{x}_2(1|\chi,T,\tau)}{\partial\tau}=-\frac{\tilde{\psi}_2(1|\chi,T,\tau)}{T}.$$

Hence,

$$\widetilde{G}_3(\chi,T|\tau)=\left(\frac{\partial\widetilde{G}_1(\chi,T|\tau)}{\partial\tau}\right)^2=\left(\frac{\tilde{\psi}_3(1|\chi,T,\tau)}{T}\right)^2,$$

and

$$\widetilde{G}_4(\chi,T|\tau)=\left(\frac{\partial\widetilde{G}_2(\chi,T|\tau)}{\partial\tau}\right)^2=\left(\frac{\tilde{\psi}_2(1|\chi,T,\tau)}{T}\right)^2.$$

□

Remark 4.1.

- To numerically solve system (4.12)-(4.13) without utilizing the explicit form of s_{delay}, let $\psi(\cdot)$ denote the corresponding state trajectory, which is defined on the original time scale. Then, it follows from (4.5) that $\psi(t_{delay})=\tilde{\psi}(s_{delay})$, in which $\tilde{\psi}(s_{delay})$ can be obtained through the application of a similar interpolation procedure, as detailed in Section 4.1 for the computation of $\tilde{x}(s_{delay})$.
- For the computation of $\dfrac{\partial x(t_{delay})}{\partial s_{delay}}$, it follows from (4.7) that

$$\frac{\partial x(t_{delay})}{\partial s_{delay}}=\frac{\partial\tilde{x}(s_{delay})}{\partial s_{delay}}=\dot{\tilde{x}}(s_{delay})=\dot{x}(t_{delay}),$$

  where $\dot{x}(t_{delay})$ can be obtained through the application of a similar procedure detailed in Section 4.1 for the computation of $x(t_{delay})$.
- For the computation of $\dfrac{\partial s_{delay}}{\partial\tau}$, taking the partial derivative of both sides of (4.2) with respect to τ gives

$$\frac{\partial\bar{\mu}(s_{delay}|T)}{\partial\tau}+\frac{\partial\bar{\mu}(s_{delay}|T)}{\partial s}\frac{\partial s_{delay}}{\partial\tau}=\frac{\partial\bar{\mu}(s|T)}{\partial\tau}=0.$$

  Then, by rearranging the equation above, we obtain

$$\frac{\partial s_{delay}}{\partial\tau}=-\frac{\partial\bar{\mu}(s_{delay}|T)}{\partial\tau}\div\frac{\partial\bar{\mu}(s_{delay}|T)}{\partial s},$$

where the values of $\frac{\partial \bar{\mu}(s_{delay}|T)}{\partial \tau}$ and $\frac{\partial \bar{\mu}(s_{delay}|T)}{\partial s}$ can also be obtained through the application of interpolation without using the explicit form of s_{delay}.

4.3. **Multi-objective optimization strategies.** In this subsection, two strategies, i.e., CWS [36] and mNBI [37], are introduced to convert $\widetilde{\mathbf{MOOP}}^{\beta}$ into a sequence of SOOCPs.

4.3.1. *Convex weighted sum.* To construct the Pareto set of $\widetilde{\mathbf{MOOP}}^{\beta}$, it is traditionally done by using CWS, which is a scalarization approach. In this scheme, $\widetilde{\mathbf{MOOP}}^{\beta}$ is transformed into a single optimization problem, through taking the convex combination of the two objective functions. That is:

$$\begin{aligned} \widetilde{\mathbf{MOOP}}^{\beta}_{\mathbf{w}}: \quad \min \quad & \widetilde{J}^{\beta}_{w}(\chi,T|\tau) := w_1\widetilde{J}^{\beta}_{1}(\chi,T|\tau) + w_2\widetilde{J}^{\beta}_{2}(\chi,T|\tau) \\ s.t. \quad & \widetilde{g}_j(\tilde{x}(s|\chi,T,\tau)) \le 0,\ \forall s\in[0,1], j\in I_{10}, \\ & (\chi,T)\in\mathscr{U}\times[a_3,b_3], \end{aligned} \tag{4.14}$$

where $w := (w_1,w_2)^{\intercal}$ is a vector of weights such that $w_1+w_2=1$ with $w_1,w_2\ge 0$.

4.3.2. *Modified normal boundary intersection.* For CWS, it has a fundamental drawback for not being able to produce an evenly spread solution points along the Pareto set, even if an equal distribution of weight vector is used. As a consequence, NBI [46] is developed to mitigate the disadvantage of CWS. A plane obtained by all convex combinations of the individual minima, i.e., the convex hull of individual minima (CHIM), is first constructed in the objective function space. Then, quasi-normal lines are constructed to this plane. The NBI scalarization scheme takes a point in CHIM and then searches for the maximum distance along the quasi-normal direction through this point. However, non-Pareto optimal solutions may be returned in some adverse cases. This means that NBI aims at getting boundary points instead of Pareto optimal solutions. Pareto solution set is a subset of boundary points. To overcome this deficiency, mNBI is formulated based on the concept of the goal attaintment approach of Gembicki [47]. A brief discription is given below. Let $(\chi_i^*,T_i^*), i=1,2$, be the respective minima of the objective functions $\widetilde{J}^{\beta}_{i}(\chi,T|\tau), i=1,2$.

$$\widetilde{\mathbf{MOOP}}^{\beta}_{\mathbf{mNBI}}: \quad \max_{\chi,T,\vartheta} \quad \widetilde{J}_{mNBI}(\vartheta) := \vartheta \tag{4.15}$$

$$\begin{aligned} s.t. \quad & \Phi w - \vartheta\Phi e \ge \widetilde{J}^{\beta}(\chi,T|\tau) \\ & \widetilde{g}_j(\tilde{x}(s|\chi,T,\tau)) \le 0,\ \forall s\in[0,1], j\in I_{10}, \\ & (\vartheta,\chi,T)\in\mathbb{R}\times\mathscr{U}\times[a_3,b_3], \end{aligned} \tag{4.16}$$

where w is a vector of weights as used in $\widetilde{\mathbf{MOOP}}^{\beta}_{\mathbf{w}}$; $e := (1,1)^{\intercal}$; Φw describes a point in CHIM; $-\Phi e$ is the normal direction at the point Φw pointing towards the origin; and

$$\Phi := \begin{pmatrix} \widetilde{J}^{\beta}_{1}(\chi_1^*,T_1^*|\tau), & \widetilde{J}^{\beta}_{1}(\chi_2^*,T_2^*|\tau) \\ \widetilde{J}^{\beta}_{2}(\chi_1^*,T_1^*|\tau), & \widetilde{J}^{\beta}_{2}(\chi_2^*,T_2^*|\tau) \end{pmatrix}^{\intercal}. \tag{4.17}$$

Compared with NBI, mNBI increases the feasible space in order to examine points which dominate the points on the CHIM. From a similar argument as that given in [37, Lemma 1], it follows that if (ϑ^*,χ^*,T^*) is an optimal solution of $\widetilde{\mathbf{MOOP}}^{\beta}_{\mathbf{mNBI}}$, then (χ^*,T^*) is weakly efficient for $\widetilde{\mathbf{MOOP}}^{\beta}$. Furthermore, a Pareto filter defined in [48] was devised to extract all non-global Pareto points. The rationale behind this filter is through a comparison of a point in the Pareto set with every other generated point. A point will be eradicated if it is not globally Pareto.

4.4. **Single-objective optimization solver.** A sequence of resulting SOOPs must be solved when using the CWS and the modified NBI strategy. In this subsection, we adopt a gradient-based single-objective optimization solver (along with the constraint transcription technique).

4.4.1. *Constraint transcription.* Continuous state inequality constraints (4.10) and (4.11) are difficult to handle directly in the numerical computation as these constraints are to be satisfied at *every* point in $[0,1]$. Two commonly used approaches to deal with these continuous state inequality constraints are: (i) the constraint transcription technique [44]; and (ii) the exact penalty method [49]. In this paper, the constraint transcription technique is used. Details are given below.

Constraints (4.10) and (4.11) can be transcribed equivalently to

$$C_j(\chi,T|\tau)=0,\ j\in I_{10}, \tag{4.18}$$

where $C_j(\chi,T|\tau):=\int_0^1 \max\left\{0,\widetilde{g}_j(\tilde{x}(s|\chi,T,\tau))\right\}\mathrm{d}s,\ j\in I_{10}$. However, $C_j(\chi,T|\tau), j\in I_{10}$, are non-smooth in $(\chi,T)\in\mathscr{U}\times[a_3,b_3]$. Accordingly, standard gradient-based optimization methods would have difficulties with these constraints. In addition, constraint (4.18) does not satisfy the linear independence constraint qualification. Thus it is not advisable to compute it as such numerically. For these reasons, we replace (4.18) with the following functional inequality constraints.

$$\widetilde{C}_j^{\varepsilon,\gamma}(\chi,T|\tau):=\int_0^1 \varphi_{\varepsilon,j}(\widetilde{g}_j(\tilde{x}(s|\chi,T,\tau)))\mathrm{d}s-\gamma>0, j\in I_{10}, \tag{4.19}$$

where $\varepsilon>0,\gamma>0$ are adjustable parameters and

$$\begin{aligned}\max\left\{0,\widetilde{g}_j(\tilde{x}(s|\chi,T,\tau))\right\}&\approx\varphi_{\varepsilon,j}(\widetilde{g}_j(\tilde{x}(s|\chi,T,\tau)))\\&:=\begin{cases}\widetilde{g}_j(\tilde{x}(s|\chi,T,\tau)),\ \text{if } \widetilde{g}_j(\tilde{x}(s|\chi,T,\tau))\in(\varepsilon,+\infty),\\ \dfrac{\left[\widetilde{g}_j(\tilde{x}(s|\chi,T,\tau))+\varepsilon\right]^2}{4\varepsilon},\ \text{if } \widetilde{g}_j(\tilde{x}(s|\chi,T,\tau))\in[-\varepsilon,\varepsilon],\\ 0,\ \text{if } \widetilde{g}_j(\tilde{x}(s|\chi,T,\tau))\in(-\infty,-\varepsilon).\end{cases}\end{aligned}$$

Remark 4.2. $\widetilde{\mathbf{MOOP}}^{\beta}_{\mathbf{mNBI}}$ and $\widetilde{\mathbf{MOOP}}^{\beta}_{\mathbf{w}}$ with constraints (4.10) and (4.11) replaced by (4.19) become standard nonlinear programming problems, referred to as $\widetilde{\mathbf{MOOP}}^{\beta,\varepsilon,\gamma}_{\mathbf{mNBI}}$ and $\widetilde{\mathbf{MOOP}}^{\beta,\varepsilon,\gamma}_{\mathbf{w}}$, respectively. Moreover, it can be shown, by similar arguments as given in [44], that, for any $\varepsilon>0$, there exists a corresponding $\gamma(\varepsilon)>0$ such that whenever $0<\gamma<\gamma(\varepsilon)$. If constraints (4.19) is satisfied, then constraints (4.10) and (4.11) are satisfied. From [44], it is noted that ε and γ are closely related to each other. At the solution of a particular problem, if a constraint is active over a large proportion of $[0,1]$, then we should choose $\gamma=O(\varepsilon)$. On the other hand, if the constraint is active only over a very small proportion of $[0,1]$, then $\gamma=O(\varepsilon^2)$ should be chosen.

4.4.2. *Gradient formulas.* To solve each of the resulting SOOPs, a sequence of standard nonlinear programming problems need to be solved by using optimization methods. Gradient-based optimization techniques, such as the SQP, work well for this purpose [44]. For this, we need the gradients of objective functions (4.14) and (4.15) as well as the gradients of constraint functions (4.16) and (4.19). For the gradients of the objective function (4.15) and the constraint function (4.16) with respect to ϑ, they can be readily calculated. For the objective function (4.14) and the constraint function (4.16), they are the functions of $\widetilde{J}^{\beta}$. Accordingly, we only need the gradients of $\widetilde{J}^{\beta}$ and constraint function $\widetilde{C}_j^{\varepsilon,\gamma}$ with respect to u and T. The derivations of these gradients are given below.

The gradients of the state $\tilde{x}(s|\chi,T,\tau)$ and $\tilde{\lambda}(s|\chi,T,\tau)$ with respect to u and T, which can be simultaneously computed via solving a set of auxiliary systems together with the governing nonlinear time-delay system, are given in Theorems 4.2-4.5 stated below.

Theorem 4.2. *For each pair* $(\chi,T)\in\mathscr{U}\times[a_3,b_3]$, $\frac{\partial\tilde{x}(s|\chi,T,\tau)}{\partial T}=\Psi(s|\chi,T,\tau)$, $s\in[0,1]$, *where* $\Psi(s|\chi,T,\tau)=\big(\Psi_1(s|\chi,T,\tau),\ldots,\Psi_5(s|\chi,T,\tau)\big)^{\mathsf{T}}\in\mathbb{R}^5$ *is the solution of the following auxiliary delay-differential system:*

$$\begin{aligned}\dot{\Psi}(s|\chi,T,\tau) &= \frac{\partial\tilde{f}(\tilde{x}(s),x(t_{delay}),T)}{\partial\tilde{x}}\Psi(s|\chi,T,\tau)+\frac{\partial\tilde{f}(\tilde{x}(s),x(t_{delay}),T)}{\partial x(t_{delay})}\Big[\Psi(s_{delay}|\chi,T,\tau)\\ &\quad+\frac{\partial x(t_{delay})}{\partial s_{delay}}\frac{\partial s_{delay}}{\partial T}\Big]+\frac{\partial\tilde{f}(\tilde{x}(s),x(t_{delay}),T)}{\partial T},s\in(0,1],\end{aligned}\tag{4.20}$$

with the initial condition

$$\Psi(s|\chi,T,\tau)=\mathbf{0}\in\mathbb{R}^5,\quad\forall s\le 0.\tag{4.21}$$

Proof. Let u and τ be arbitrary but fixed, and let $\mathbf{e}$ be the unit vector in $\mathbb{R}$. Then,

$$\frac{\partial\tilde{x}(s)}{\partial T}=\lim_{\xi\to0}\frac{\tilde{x}(s|\chi,T^{\xi},\tau)-\tilde{x}(s|\chi,T,\tau)}{\xi},\tag{4.22}$$

where $T^{\xi}=T+\xi\mathbf{e}$. s^{ξ}_{delay} and t^{ξ}_{delay} are the corresponding delayed time points in $[0,1]$ and $[0,T]$ such that $\bar{\mu}(s^{\xi}_{delay}|T^{\xi})=t^{\xi}_{delay}=\bar{\mu}(s|T^{\xi})-\tau,s\in[0,1]$. To proceed, we will prove this theorem in four steps.

Step 1. Define $F^{\xi}(\varsigma):=\tilde{f}(\tilde{x}(\varsigma),\tilde{x}(\varsigma^{\xi}_{delay}),T^{\xi})$. For each real number $\xi\in\mathbb{R}$, let $\tilde{x}^{\xi}(\cdot)$ denote the function $\tilde{x}(\cdot|\chi,T^{\xi},\tau)$. Then, from (4.4), for each $\xi\in\mathbb{R}$, we have

$$\begin{aligned}\tilde{x}^{\xi}(s)&=\tilde{x}(0)+\int_0^s F^{\xi}(\varsigma)\mathrm{d}\varsigma\\ &=\tilde{x}(0)+\int_0^s\tilde{f}(\tilde{x}(\varsigma),\tilde{x}(\varsigma^{\xi}_{delay}),T^{\xi})\mathrm{d}\varsigma\\ &=\tilde{x}(0)+\int_0^s T^{\xi}f(\tilde{x}(\varsigma),\tilde{x}(\varsigma^{\xi}_{delay}))\mathrm{d}\varsigma,s\in[0,1],\end{aligned}$$

where $\tilde{x}(\cdot)$ denote the function $\tilde{x}(\cdot|\chi,T,\tau)$. Define

$$\Gamma^{\xi}(s)=\tilde{x}(s|\chi,T^{\xi},\tau)-\tilde{x}(s|\chi,T,\tau)=\int_0^s(F^{\xi}(\varsigma)-F(\varsigma))\mathrm{d}\varsigma.\tag{4.23}$$

It follows from the mean value theorem that

$$F^{\xi}(\varsigma)-F(\varsigma) = \int_0^1\left\{\frac{\partial\tilde{f}^{\xi}_{\eta}}{\partial\tilde{x}}\Gamma^{\xi}(s)+\frac{\partial\tilde{f}^{\xi}_{\eta}}{\partial\tilde{x}_d}(\tilde{x}^{\xi}_d-\tilde{x}_d)+\frac{\partial\tilde{f}^{\xi}_{\eta}}{\partial T}\xi\right\}\mathrm{d}\eta,s\in[0,1],\tag{4.24}$$

where $\tilde{f}^{\xi}_{\eta}:=\tilde{f}(\tilde{x}(s)+\eta\Gamma^{\xi}(s),\tilde{x}_d+\eta(\tilde{x}^{\xi}_d-\tilde{x}_d),T+\eta\xi\mathbf{e})$ and

$$\begin{aligned}\tilde{x}^{\xi}_d-\tilde{x}_d &= \tilde{x}(s^{\xi}_{delay}|\chi,T^{\xi},\tau)-\tilde{x}(s_{delay}|\chi,T,\tau)\\ &= \tilde{x}(s^{\xi}_{delay}|\chi,T^{\xi},\tau)-\tilde{x}(s_{delay}|\chi,T^{\xi},\tau)+\tilde{x}(s_{delay}|\chi,T^{\xi},\tau)-\tilde{x}(s_{delay}|\chi,T,\tau)\\ &= \tilde{x}(s^{\xi}_{delay}|\chi,T^{\xi},\tau)-\tilde{x}(s_{delay}|\chi,T^{\xi},\tau)+\Gamma^{\xi}(s_{delay}).\end{aligned}\tag{4.25}$$

The state set $\{\tilde{x}^{\xi}(s):\xi\in[-1,1]\}$ is equi-bounded on $[0,1]$ due to the fact that $\tilde{f}(\tilde{x}(s),x(t_{delay}),T)$ satisfies the linear growth condition. Accordingly, there exists a real number $c_1>0$ such that, for each $\xi\in[-1,1]$, $\tilde{x}^{\xi}(s)\in\mathscr{N}_5(c_1)$, where $\mathscr{N}_5(c_1)$ denotes the closed ball in $\mathbb{R}^5$ of radius c_1 centered at the origin. It should be noted that $\mathscr{N}_5(c_1)$ is convex, therefore, for each $\xi\in[-1,1]$, $\tilde{x}(s)+\eta\Gamma^{\xi}(s)\in\mathscr{N}_5(c_1)$, $s\in[0,1]$, $\eta\in[0,1]$. Moreover, it is clear that, for each $\xi\in[-1,1]$, $T+\eta\xi\mathbf{e}\in\mathscr{N}_1(c_2)$, where $c_2:=|T|+1$. From (2.1) and (4.4), it follows that $\dfrac{\partial\tilde{f}}{\partial\tilde{x}}$ and $\dfrac{\partial\tilde{f}}{\partial\tilde{x}_d}$ are continuous. Consequently, by the compactness of $[0,1]$,

$\mathcal{N}_5(c_1)$ and $\mathcal{N}_1(c_2)$ together with the definitions of ϕ, there exists a real number $c_3 > 0$ such that, for each $\xi \in [-1,1]$,

$$\left|\frac{\partial \tilde{f}_\eta^\xi}{\partial \tilde{x}}\right|_{5\times 5} \le c_3,\ \left|\frac{\partial \tilde{f}_\eta^\xi}{\partial \tilde{x}_d}\right|_{5\times 5} \le c_3,\ \left|\frac{\partial \tilde{f}_\eta^\xi}{\partial T}\right|_{5} \le c_3,\ \left|\frac{\partial \phi_\eta^\xi}{\partial \bar{\mu}}\right|_{5} \le c_3,\ s \in [0,1], \eta \in [0,1], \tag{4.26}$$

where $\phi_\eta^\xi := \phi(\bar{\mu}(s|T+\eta\xi\mathbf{e}) - \tau)$, and $\left|\cdot\right|$ denotes the Euclidian norm.

Step 2. The function $\Gamma^\xi(s)$ is of order ξ. Let $\xi \in [-1,1]$ be arbitrary. Let s_p denote a time point in the new time horizon such that $\bar{\mu}(s_p|T) = \tau$. It is clear from inequalities (4.23) and (4.24) that

$$\begin{aligned}
&|\Gamma^\xi(s)|_5 \\
&= \left|\int_0^s\int_0^1 \left\{\frac{\partial \tilde{f}_\eta^\xi}{\partial \tilde{x}}\Gamma^\xi(s) + \frac{\partial \tilde{f}_\eta^\xi}{\partial \tilde{x}_d}\left[\Gamma^\xi(s_{delay}) + \tilde{x}(s_{delay}^\xi|\chi,T^\xi,\tau) - \tilde{x}(s_{delay}|\chi,T^\xi,\tau)\right] + \frac{\partial \tilde{f}_\eta^\xi}{\partial T}\xi\right\}\mathrm{d}\eta\mathrm{d}\varsigma\right|_5 \\
&\le \left|\int_0^s\int_0^1 \frac{\partial \tilde{f}_\eta^\xi}{\partial \tilde{x}}\Gamma^\xi(s)\mathrm{d}\eta\mathrm{d}\varsigma\right|_5 + \left|\int_0^s\int_0^1 \frac{\partial \tilde{f}_\eta^\xi}{\partial \tilde{x}_d}\Gamma^\xi(s_{delay})\mathrm{d}\eta\mathrm{d}\varsigma\right|_5 \\
&\quad + \left|\int_0^s\int_0^1 \left[\tilde{x}(s_{delay}^\xi|\chi,T^\xi,\tau) - \tilde{x}(s_{delay}|\chi,T^\xi,\tau)\right]\mathrm{d}\eta\mathrm{d}\varsigma\right|_5 + \left|\int_0^s\int_0^1 \frac{\partial \tilde{f}_\eta^\xi}{\partial T}\xi\mathrm{d}\eta\mathrm{d}\varsigma\right|_5 \\
&= \left|\int_0^{s_p}\int_0^1 \frac{\partial \tilde{f}_\eta^\xi}{\partial \tilde{x}}\Gamma^\xi(s)\mathrm{d}\eta\mathrm{d}\varsigma\right|_5 + \left|\int_0^{s_p}\int_0^1 \frac{\partial \tilde{f}_\eta^\xi}{\partial \tilde{x}_d}\Gamma^\xi(s_{delay})\mathrm{d}\eta\mathrm{d}\varsigma\right|_5 \\
&\quad + \left|\int_0^{s_p}\int_0^1 \left[\tilde{x}(s_{delay}^\xi|\chi,T^\xi,\tau) - \tilde{x}(s_{delay}|\chi,T^\xi,\tau)\right]\mathrm{d}\eta\mathrm{d}\varsigma\right|_5 + \left|\int_0^{s_p}\int_0^1 \frac{\partial \tilde{f}_\eta^\xi}{\partial T}\xi\mathrm{d}\eta\mathrm{d}\varsigma\right|_5 \\
&\quad + \left|\int_{s_p}^s\int_0^1 \frac{\partial \tilde{f}_\eta^\xi}{\partial \tilde{x}}\Gamma^\xi(s)\mathrm{d}\eta\mathrm{d}\varsigma\right|_5 + \left|\int_{s_p}^s\int_0^1 \frac{\partial \tilde{f}_\eta^\xi}{\partial \tilde{x}_d}\Gamma^\xi(s_{delay})\mathrm{d}\eta\mathrm{d}\varsigma\right|_5 \\
&\quad + \left|\int_{s_p}^s\int_0^1 \left[\tilde{x}(s_{delay}^\xi|\chi,T^\xi,\tau) - \tilde{x}(s_{delay}|\chi,T^\xi,\tau)\right]\mathrm{d}\eta\mathrm{d}\varsigma\right|_5 + \left|\int_{s_p}^s\int_0^1 \frac{\partial \tilde{f}_\eta^\xi}{\partial T}\xi\mathrm{d}\eta\mathrm{d}\varsigma\right|_5, s \in [0,1].
\end{aligned}$$

Case 1. When $s_{delay} < 0$ namely, $s \in [0, s_p]$, taking the norm of both sides of (4.23) and applying the definition of c_3 yields

$$|\Gamma^\xi(s)|_5 = \left|\int_0^s\int_0^1 \left\{\frac{\partial \tilde{f}_\eta^\xi}{\partial \tilde{x}}\Gamma^\xi(s) + \frac{\partial \tilde{f}_\eta^\xi}{\partial \tilde{x}_d}(\phi_\eta^\xi - \phi_\eta) + \frac{\partial \tilde{f}_\eta^\xi}{\partial T}\xi\right\}\mathrm{d}\eta\mathrm{d}\varsigma\right|_5,$$

where

$$\phi_\eta^\xi - \phi_\eta = \phi(\bar{\mu}(s|T+\xi\mathbf{e}) - \tau) - \phi(\bar{\mu}(s|T) - \tau) = \xi\frac{\partial\phi(\bar{\mu}(s|T+\eta\xi\mathbf{e}) - \tau)}{\partial\bar{\mu}}\frac{\partial\bar{\mu}(s|T+\eta\xi\mathbf{e})}{\partial T}, \eta \in [0,1].$$

It follows that

$$|\phi_\eta^\xi - \phi_\eta| \le \xi\left|\frac{\partial\phi(\bar{\mu}(s|T+\eta\xi\mathbf{e}) - \tau)}{\partial\bar{\mu}}\right|\left|\frac{\partial\bar{\mu}(s|T+\eta\xi\mathbf{e})}{\partial T}\right| \le |\xi|c_3T, \eta \in [0,1]. \tag{4.27}$$

and then

$$|\Gamma^\xi(s)|_5 \le c_3|\xi| + c_3^2T|\xi| + \int_0^s c_3|\Gamma^\xi(\varsigma)|_5\mathrm{d}\varsigma,\ s \in [0, s_p].$$

It follows from Gronwall's Lemma (see [50, Theorem 2.8.6]) that

$$|\Gamma^{\xi}(s)|_5 \leq |\xi|(c_3+c_3^2T)\exp(c_3),\ s\in[0,s_p]. \tag{4.28}$$

Case 2. When $s_{delay}\geq 0$, namely, $s\in[s_p,1]$, one has

$$\begin{aligned}
&|\Gamma^{\xi}(s)|_5\\
&=\left|\int_0^s\int_0^1\left\{\frac{\partial\tilde{f}_\eta^\xi}{\partial\tilde{x}}\Gamma^{\xi}(s)+\frac{\partial\tilde{f}_\eta^\xi}{\partial\tilde{x}_d}\Big[\Gamma^{\xi}(s_{delay})+\tilde{x}(s_{delay}^{\xi}|\chi,T^{\xi},\tau)-\tilde{x}(s_{delay}|\chi,T^{\xi},\tau)\Big]+\frac{\partial\tilde{f}_\eta^\xi}{\partial T}\xi\right\}\mathrm{d}\eta\mathrm{d}\varsigma\right|_5\\
&\leq|\xi|(c_3+c_3^2T)\exp(c_3)+\left|\int_{s_p}^s\int_0^1\frac{\partial\tilde{f}_\eta^\xi}{\partial\tilde{x}}\Gamma^{\xi}(s)\mathrm{d}\eta\mathrm{d}\varsigma\right|_5+\left|\int_{s_p}^s\int_0^1\frac{\partial\tilde{f}_\eta^\xi}{\partial\tilde{x}_d}\Gamma^{\xi}(s_{delay})\mathrm{d}\eta\mathrm{d}\varsigma\right|_5\\
&\quad+\left|\int_{s_p}^s\int_0^1\Big[\tilde{x}(s_{delay}^{\xi}|\chi,T^{\xi},\tau)-\tilde{x}(s_{delay}|\chi,T^{\xi},\tau)\Big]\mathrm{d}\eta\mathrm{d}\varsigma\right|_5+\left|\int_{s_p}^s\int_0^1\frac{\partial\tilde{f}_\eta^\xi}{\partial T}\xi\mathrm{d}\eta\mathrm{d}\varsigma\right|_5.
\end{aligned}$$

Applying the definition of s_{delay} and using the fact that $s_{delay}>0$, we obtain

$$\left|\int_{s_p}^s\int_0^1\frac{\partial\tilde{f}_\eta^\xi}{\partial\tilde{x}_d}\Gamma^{\xi}(s_{delay})\mathrm{d}\eta\mathrm{d}\varsigma\right|_5\leq\left|\int_{s_p}^s c_3|\Gamma^{\xi}(s_{delay})|_5\mathrm{d}\varsigma\right|_5\leq\left|\int_0^s c_3|\Gamma^{\xi}(\varsigma)|_5\mathrm{d}\varsigma\right|_5.$$

It follows from the mean value theorem that

$$|\tilde{x}(s_{delay}^{\xi}|\chi,T^{\xi},\tau)-\tilde{x}(s_{delay}|\chi,T^{\xi},\tau)|_5\leq\left|\frac{\partial\tilde{x}(z(s|T+\eta\xi\mathbf{e})|\chi,T+\eta\xi\mathbf{e},\tau)}{\partial z}\right|\left|\frac{\partial z}{\partial T}\right||\xi|\leq c_3|\xi|, \tag{4.29}$$

where $z(s|T+\eta\xi\mathbf{e})=s_{delay}+\eta(s_{delay}^{\xi}-s_{delay}),\eta\in[0,1]$. By Gronwall's Lemma (see [50, Theorem 2.8.6]), we have

$$\begin{aligned}
|\Gamma^{\xi}(s)|_5&\leq|\xi|(c_3+c_3^2T)\exp(c_3)+c_3|\xi|+c_3^2|\xi|+\int_0^s 2c_3|\Gamma^{\xi}(\tau)|_5\mathrm{d}\varsigma\\
&\leq|\xi|\Big[c_3+c_3^2+(c_3+c_3^2T)\exp(c_3)\Big]\exp(2c_3),s\in[s_p,1].
\end{aligned}\tag{4.30}$$

It follows from (4.28) and (4.30) that

$$|\Gamma^{\xi}(s)|_5\leq|\xi|\Big[c_3+c_3^2+(c_3+c_3^2T)\exp(c_3)\Big]\exp(2c_3),s\in[0,1]. \tag{4.31}$$

Since $\xi\in[-1,1]$ is arbitrary, we have that $\Gamma^{\xi}(s)$ is of order ξ. In view of (4.25)-(4.27), (4.29), and (4.31), we have

$$\begin{aligned}
|\tilde{x}_d^{\xi}-\tilde{x}_d|_5&\leq|\tilde{x}(s_{delay}^{\xi}|\chi,T^{\xi},\tau)-\tilde{x}(s_{delay}|\chi,T^{\xi},\tau)|_5+|\Gamma^{\xi}(s_{delay})|_5\\
&\leq c_3|\xi|(1+T)+|\xi|\Big[c_3+c_3^2+(c_3+c_3^2T)\exp(c_3)\Big]\exp(2c_3).
\end{aligned}\tag{4.32}$$

Step 3. The definition of ρ and its properties.

Let the function $\rho:[-1,0)\cup(0,1]$ be defined by

$$\rho(\xi)=|\xi|^{-1}\int_0^1\Big\{|\lambda^{1,\xi}(s)|_5+|\lambda^{2,\xi}(s)|_5+|\lambda^{3,\xi}(s)|_5\Big\}\mathrm{d}s, \tag{4.33}$$

where

$$\lambda^{1,\xi}(s) = \int_0^1 \left\{ \frac{\partial \tilde{f}(\tilde{x}+\eta\Gamma^{\xi}(s),\tilde{x}_d+\eta(\tilde{x}_d^{\xi}-\tilde{x}_d),T+\eta\xi\mathbf{e})}{\partial \tilde{x}} - \frac{\partial \tilde{f}(\tilde{x},\tilde{x}_d,T)}{\partial \tilde{x}} \right\} \Gamma^{\xi}(s)\mathrm{d}\eta, \quad (4.34)$$

$$\lambda^{2,\xi}(s) = \int_0^1 \left\{ \frac{\partial \tilde{f}(\tilde{x}+\eta\Gamma^{\xi}(s),\tilde{x}_d+\eta(\tilde{x}_d^{\xi}-\tilde{x}_d),T+\eta\xi\mathbf{e})}{\partial \tilde{x}_d} - \frac{\partial \tilde{f}(\tilde{x},\tilde{x}_d,T)}{\partial \tilde{x}_d} \right\} (\tilde{x}_d^{\xi}-\tilde{x}_d)\mathrm{d}\eta, \quad (4.35)$$

$$\lambda^{3,\xi}(s) = \int_0^1 \left\{ \frac{\partial \tilde{f}(\tilde{x}+\eta\Gamma^{\xi}(s),\tilde{x}_d+\eta(\tilde{x}_d^{\xi}-\tilde{x}_d),T+\eta\xi\mathbf{e})}{\partial T} - \frac{\partial \tilde{f}(\tilde{x},\tilde{x}_d,T)}{\partial T} \right\} \xi\mathrm{d}\eta. \quad (4.36)$$

By (4.26), (4.31), and (4.32), we obtain

$$|\xi|^{-1}|\lambda^{1,\xi}(s)| \leq 2c_3\left(\left[(c_3+c_3^2T)\exp(c_3)+(c_3+3c_3^2)\right]\exp(2c_3)\right), \quad (4.37)$$

$$|\xi|^{-1}|\lambda^{2,\xi}(s)| \leq 2c_3\left(c_3(1+T)+\left[(c_3+c_3^2T)\exp(c_3)+(c_3+3c_3^2)\right]\exp(2c_3)\right), \quad (4.38)$$

$$|\xi|^{-1}|\lambda^{3,\xi}(s)| \leq 2c_3. \quad (4.39)$$

From (4.31)-(4.32), we conclude

$$\tilde{x}+\eta\Gamma^{\xi}(s) \to \tilde{x}, \text{ as } \xi\to 0, \quad (4.40)$$

$$\tilde{x}_d+\eta(\tilde{x}_d^{\xi}-\tilde{x}_d) \to \tilde{x}_d, \text{ as } \xi\to 0, \quad (4.41)$$

uniformly with respect to $s\in[0,1]$ and $\eta\in[0,1]$. Furthermore, we have

$$T+\eta\xi\mathbf{e} \to T, \text{ as } \xi\to 0, \quad (4.42)$$

uniformly with respect to $\eta\in[0,1]$. Since the convergence in (4.40) and (4.41) takes place inside the ball $\mathscr{N}_5(c_1)$, and the convergence in (4.42) takes place inside of the ball $\mathscr{N}_1(c_2)$ together with the fact that $\frac{\partial \tilde{f}}{\partial \tilde{x}}$, $\frac{\partial \tilde{f}}{\partial \tilde{x}_d}$ and $\frac{\partial \tilde{f}}{\partial T}$ are uniformly continuous on the compact set $[0,1]\times\mathscr{N}_5(c_1)\times\mathscr{N}_5(c_1)\times\mathscr{N}_1(c_2)$, we can show that

$$\begin{aligned}
\frac{\partial \tilde{f}(\tilde{x}+\eta\Gamma^{\xi}(s),\tilde{x}_d+\eta(\tilde{x}_d^{\xi}-\tilde{x}_d),T+\eta\xi\mathbf{e})}{\partial \tilde{x}} &\to \frac{\partial \tilde{f}(\tilde{x}(s),\tilde{x}_d,T)}{\partial \tilde{x}}, \text{ as } \xi\to 0,\\
\frac{\partial \tilde{f}(\tilde{x}(s)+\eta\Gamma^{\xi}(s),\tilde{x}_d+\eta(\tilde{x}_d^{\xi}-\tilde{x}_d),T+\eta\xi\mathbf{e})}{\partial \tilde{x}_d} &\to \frac{\partial \tilde{f}(\tilde{x}(s),\tilde{x}_d,T)}{\partial \tilde{x}_d}, \text{ as } \xi\to 0,\\
\frac{\partial \tilde{f}(\tilde{x}(s)+\eta\Gamma^{\xi}(s),\tilde{x}_d+\eta(\tilde{x}_d^{\xi}-\tilde{x}_d),T+\eta\xi\mathbf{e})}{\partial T} &\to \frac{\partial \tilde{f}(\tilde{x}(s),\tilde{x}_d,T)}{\partial T}, \text{ as } \xi\to 0,\\
\frac{\partial \tilde{x}(s_{delay}^{\xi,\eta}|\chi,T+\eta\xi\mathbf{e},\tau)}{\partial s_{delay}} &\to \frac{\partial \tilde{x}(s_{delay}|\chi,T,\tau)}{\partial s_{delay}}, \text{ as } \xi\to 0,\\
\frac{\partial s_{delay}^{\xi,\eta}}{\partial T} &\to \frac{\partial s_{delay}}{\partial T}, \text{ as } \xi\to 0,
\end{aligned}$$

uniformly with respect to $s\in[0,1]$ and $\eta\in[0,1]$, where $s_{delay}^{\xi,\eta} := s_{delay}+\eta(s_{delay}^{\xi}-s_{delay})$ is the corresponding delayed time $T+\eta\xi\mathbf{e}$. These results along with (4.31) imply that

$$|\xi|^{-1}|\lambda^{1,\xi}(s)|\to 0, |\xi|^{-1}|\lambda^{2,\xi}(s)|\to 0, |\xi|^{-1}|\lambda^{3,\xi}(s)|\to 0, \text{ as } \xi\to 0, \quad (4.43)$$

uniformly with respect to $s\in[0,1]$. Based on (4.37)-(4.39) and (4.43), the conditions of Lebesgue Dominated convergence theorem (see Theorem 2.6.4 in [50]) hold. Therefore,

$$\lim_{\xi\to 0}\rho(\xi)=\int_0^1\lim_{\xi\to 0}\left\{|\xi|^{-1}|\lambda^{1,\xi}(s)|_8+|\xi|^{-1}|\lambda^{2,\xi}(s)|_8+|\xi|^{-1}|\lambda^{3,\xi}(s)|_8\right\}\mathrm{d}s=0. \tag{4.44}$$

Step 4. Comparing $\xi^{-1}\Gamma^{\xi}$ with $\Psi(\cdot|\chi,T,\tau)$.

Let $\xi\in[-1,0)\cup(0,1]$ be arbitrary but fixed. Next, it follows from (4.23)-(4.25), (4.34)-(4.36) and the mean value theorem that

$$\begin{aligned}
\Gamma^{\xi}(s) &= \tilde{x}(s|\chi,T^{\xi},\tau)-\tilde{x}(s|\chi,T,\tau)=\int_0^s\int_0^1\frac{\partial\tilde{f}^{\xi}_{\eta}}{\partial\tilde{x}}\Gamma^{\xi}(s)+\frac{\partial\tilde{f}^{\xi}_{\eta}}{\partial\tilde{x}_d}(\tilde{x}^{\xi}_d-\tilde{x}_d)+\frac{\partial\tilde{f}^{\xi}_{\eta}}{\partial T}\xi\mathrm{d}\eta\mathrm{d}\varsigma\\
&= \int_0^s\left\{\sum_{i=1}^3\lambda^{i,\xi}(\varsigma)\right\}\mathrm{d}\varsigma+\int_0^s\frac{\partial\tilde{f}(\tilde{x},\tilde{x}_d,T)}{\partial\tilde{x}}\Gamma^{\xi}(\varsigma)\mathrm{d}\varsigma\\
&\quad+\int_0^s\frac{\partial\tilde{f}(\tilde{x},\tilde{x}_d,T)}{\partial\tilde{x}_d}(\tilde{x}^{\xi}_d-\tilde{x}_d)\mathrm{d}\varsigma+\int_0^s\frac{\partial\tilde{f}(\tilde{x},\tilde{x}_d,T)}{\partial T}\xi\mathrm{d}\varsigma\\
&= \int_0^s\left\{\sum_{i=1}^3\lambda^{i,\xi}(\varsigma)\right\}\mathrm{d}\varsigma+\int_0^s\frac{\partial\tilde{f}(\tilde{x},\tilde{x}_d,T)}{\partial\tilde{x}}\Gamma^{\xi}(\varsigma)\mathrm{d}\varsigma+\int_0^s\frac{\partial\tilde{f}(\tilde{x},\tilde{x}_d,T)}{\partial T}\xi\mathrm{d}\varsigma\\
&\quad+\int_0^s\frac{\partial\tilde{f}(\tilde{x},\tilde{x}_d,T)}{\partial\tilde{x}_d}\left[\Gamma^{\xi}(s_{delay})+\tilde{x}(s^{\xi}_{delay}|\chi,T^{\xi},\tau)-\tilde{x}(s_{delay}|\chi,T^{\xi},\tau)\right]\mathrm{d}\varsigma\\
&= \int_0^s\left\{\sum_{i=1}^3\lambda^{i,\xi}(\varsigma)\right\}\mathrm{d}\varsigma+\int_0^s\frac{\partial\tilde{f}(\tilde{x},\tilde{x}_d,T)}{\partial\tilde{x}}\Gamma^{\xi}(\varsigma)\mathrm{d}\varsigma+\int_0^s\frac{\partial\tilde{f}(\tilde{x},\tilde{x}_d,T)}{\partial T}\xi\mathrm{d}\varsigma\\
&\quad+\int_0^s\frac{\partial\tilde{f}(\tilde{x},\tilde{x}_d,T)}{\partial\tilde{x}_d}\left[\Gamma^{\xi}(s_{delay})+\frac{\partial\tilde{x}(s^{\xi,\eta}_{delay}|\chi,T+\eta\xi\mathbf{e},\tau)}{\partial s_{delay}}\frac{\partial s^{\xi,\eta}_{delay}}{\partial T}\xi\right]\mathrm{d}\varsigma.
\end{aligned}\tag{4.45}$$

Observe $\tilde{x}_d=\tilde{x}(s_{delay}):=\tilde{x}(s_{delay}|\chi,T,\tau)=x(t_{delay}):=x(t_{delay}|\chi,T,\tau)$. Taking the integration of auxiliary system (4.20), we conclude that

$$\begin{aligned}
\Psi(s|\chi,T,\tau) &= \int_0^s\frac{\partial\tilde{f}(\tilde{x},x(t_{delay}),T)}{\partial\tilde{x}}\Psi(\varsigma|\chi,T,\tau)\mathrm{d}\varsigma+\int_0^s\frac{\partial\tilde{f}(\tilde{x},x(t_{delay}),T)}{\partial T}\mathrm{d}\varsigma,\\
&\quad+\int_0^s\frac{\partial\tilde{f}(\tilde{x},x(t_{delay}),T)}{\partial x(t_{delay})}\left[\Psi(s_{delay}|\chi,T,\tau)+\frac{\partial x(t_{delay})}{\partial s_{delay}}\frac{\partial s_{delay}}{\partial T}\right]\mathrm{d}\varsigma,\\
&= \int_0^s\frac{\partial\tilde{f}(\tilde{x},\tilde{x}_d,T)}{\partial\tilde{x}}\Psi(\varsigma|\chi,T,\tau)\mathrm{d}\varsigma+\int_0^s\frac{\partial\tilde{f}(\tilde{x},\tilde{x}_d,T)}{\partial T}\mathrm{d}\varsigma,\\
&\quad+\int_0^s\frac{\partial\tilde{f}(\tilde{x},\tilde{x}_d,T)}{\partial\tilde{x}_d}\left[\Psi(s_{delay}|\chi,T,\tau)+\frac{\partial\tilde{x}(s_{delay}|\chi,T,\tau)}{\partial s_{delay}}\frac{\partial s_{delay}}{\partial T}\right]\mathrm{d}\varsigma.
\end{aligned}\tag{4.46}$$

Multiplying (4.45) by ξ^{-1} and subtracting it from (4.46), we obtain

$$\begin{aligned}
&\xi^{-1}\Gamma^{\xi}(s)-\Psi(s|\chi,T,\tau)\\
&=\xi^{-1}\int_0^s\left\{\sum_{i=1}^{3}\lambda^{i,\xi}(\varsigma)\right\}\mathrm{d}\varsigma+\int_0^s\frac{\partial\tilde{f}(\tilde{x},x(t_{delay}),T)}{\partial\tilde{x}}\left[\xi^{-1}\Gamma^{\xi}(\varsigma)-\Psi(\varsigma|\chi,T,\tau)\right]\mathrm{d}\varsigma\\
&\quad+\int_0^s\frac{\partial\tilde{f}(\tilde{x},\tilde{x}_d,T)}{\partial\tilde{x}_d}\left[\left(\xi^{-1}\Gamma^{\xi}(s_{delay})-\Psi(s_{delay}|\chi,T,\tau)\right)\right.\\
&\quad\left.+\frac{\partial\tilde{x}(s^{\xi,\eta}_{delay}|\chi,T+\eta\xi\mathbf{e},\tau)}{\partial s_{delay}}\frac{\partial s^{\xi,\eta}_{delay}}{\partial T}-\frac{\partial\tilde{x}(s_{delay}|\chi,T,\tau)}{\partial s_{delay}}\frac{\partial s_{delay}}{\partial T}\right]\mathrm{d}\varsigma.
\end{aligned}$$

In view of the definition of $\rho(\xi)$ in (4.33), we have

$$\begin{aligned}
&|\xi^{-1}\Gamma^{\xi}(s)-\Psi(s|\chi,T,\tau)|_5\\
&=\left|\xi^{-1}\int_0^s\left\{\sum_{i=1}^{3}\lambda^{i,\xi}(\varsigma)\right\}\mathrm{d}\varsigma+\int_0^s\frac{\partial\tilde{f}(\tilde{x},x(t_{delay}),T)}{\partial\tilde{x}}\left[\xi^{-1}\Gamma^{\xi}(\varsigma)-\Psi(\varsigma|\chi,T,\tau)\right]\mathrm{d}\varsigma\right.\\
&\quad+\int_0^s\frac{\partial\tilde{f}(\tilde{x},\tilde{x}_d,T)}{\partial\tilde{x}_d}\left[\left(\xi^{-1}\Gamma^{\xi}(s_{delay})-\Psi(s_{delay}|\chi,T,\tau)\right)\right.\\
&\quad\left.\left.+\frac{\partial\tilde{x}(s^{\xi,\eta}_{delay}|\chi,T+\eta\xi\mathbf{e},\tau)}{\partial s_{delay}}\frac{\partial s^{\xi,\eta}_{delay}}{\partial T}-\frac{\partial\tilde{x}(s_{delay}|\chi,T,\tau)}{\partial s_{delay}}\frac{\partial s_{delay}}{\partial T}\right]\mathrm{d}\varsigma\right|\\
&\le\rho(\xi)+\int_0^s c_3\left|\xi^{-1}\Gamma^{\xi}(\varsigma)-\Psi(\varsigma|\chi,T,\tau)\right|_5\mathrm{d}\varsigma+\int_0^s c_3\left|\xi^{-1}\Gamma^{\xi}(s_{delay})-\Psi(s_{delay}|\chi,T,\tau)\right|_5\mathrm{d}\varsigma\\
&\quad+\int_0^s c_3\left|\frac{\partial\tilde{x}(s^{\xi,\eta}_{delay}|\chi,T+\eta\xi\mathbf{e},\tau)}{\partial s_{delay}}\frac{\partial s^{\xi,\eta}_{delay}}{\partial T}-\frac{\partial\tilde{x}(s_{delay}|\chi,T,\tau)}{\partial s_{delay}}\frac{\partial s_{delay}}{\partial T}\right|_5\mathrm{d}\varsigma\\
&\le\tilde{\rho}(\xi)+\int_0^s 2c_3\left|\xi^{-1}\Gamma^{\xi}(\varsigma)-\Psi(\varsigma|\chi,T,\tau)\right|_5\mathrm{d}\varsigma,
\end{aligned}$$

where

$$\tilde{\rho}(\xi):=\rho(\xi)+\int_0^s c_3\left|\frac{\partial\tilde{x}(s^{\xi,\eta}_{delay}|\chi,T+\eta\xi\mathbf{e},\tau)}{\partial s_{delay}}\frac{\partial s^{\xi,\eta}_{delay}}{\partial T}-\frac{\partial\tilde{x}(s_{delay}|\chi,T,\tau)}{\partial s_{delay}}\frac{\partial s_{delay}}{\partial T}\right|_5\mathrm{d}\varsigma.$$

Clearly, $\tilde{\rho}(\xi)\to 0$ as $\xi\to 0$. It follows from Gronwall's Lemma (see [50, Theorem 2.8.6]) that

$$|\xi^{-1}\Gamma^{\xi}(s)-\Psi(s|\chi,T,\tau)|_5\le\tilde{\rho}(\xi)\exp(2c_3),s\in[0,1]. \tag{4.47}$$

Note that $\xi\in[-1,0)\cup(0,1]$ is arbitrary. Therefore, by taking the limit as $\xi\to 0$ in (4.47), it follows from (4.44) that

$$\lim_{\xi\to 0}\xi^{-1}\Gamma^{\xi}(s)=\Psi(s|\chi,T,\tau),s\in[0,1]. \tag{4.48}$$

From (4.22), (4.23), and (4.48), we obtain

$$\frac{\partial\tilde{x}(s|\chi,T,\tau)}{\partial T}=\lim_{\xi\to 0}\xi^{-1}[\tilde{x}(s|\chi,T^{\xi},\tau)-\tilde{x}(s|\chi,T,\tau)]=\lim_{\xi\to 0}\xi^{-1}\Gamma^{\xi}(s)=\Psi(s|\chi,T,\tau),s\in[0,1].$$

□

Theorem 4.3. *For each pair* $(\chi,T)\in\mathscr{U}\times[a_3,b_3]$, $\frac{\partial\tilde{x}(s|\chi,T,\tau)}{\partial u}=\Theta(s|\chi,T,\tau)$, $s\in[0,1]$, *where* $\Theta(s|\chi,T,\tau)=\big(\Theta_{ij}(s|\chi,T,\tau)\big)_{5\times2}\in\mathbb{R}^{5\times2}$ *is the solution of the following auxiliary delay-differential system:*

$$\dot{\Theta}(s|\chi,T,\tau)=\frac{\partial\tilde{f}(\tilde{x}(s),x(t_{delay}),T)}{\partial\tilde{x}}\Theta(s|\chi,T,\tau)+\frac{\partial\tilde{f}(\tilde{x}(s),x(t_{delay}),T)}{\partial x(t_{delay})}\Theta(s_{delay}|\chi,T,\tau),s\in(0,1], \tag{4.49}$$

with the initial condition

$$\Theta(s|\chi,T,\tau)=\mathbf{0}\in\mathbb{R}^{5\times2},\quad\forall s\le0. \tag{4.50}$$

Proof. From Theorem 4.2, we can conclude the desired conclusion immediately. □

Define

$$\begin{aligned}&\tilde{h}(\tilde{x}(s),x(t_{delay}),T)\\&:=\frac{\partial\tilde{f}(\tilde{x}(s),x(t_{delay}),T)}{\partial\tilde{x}}\tilde{\lambda}(s|\chi,T,\tau)+\frac{\partial\tilde{f}(\tilde{x}(s),x(t_{delay}),T)}{\partial x(t_{delay})}\Big[\tilde{\lambda}(s_{delay}|\chi,T,\tau)+\frac{\partial x(t_{delay})}{\partial s_{delay}}\frac{\partial s_{delay}}{\partial\tau}\Big].\end{aligned}$$

Theorem 4.4. *For each pair* $(\chi,T)\in\mathscr{U}\times[a_3,b_3]$, $\frac{\partial\tilde{\Psi}(s|\chi,T,\tau)}{\partial T}=\Upsilon(s|\chi,T,\tau)$, $s\in[0,1]$, *where* $\Upsilon(s|\chi,T,\tau)=\big(\Upsilon_1(s|\chi,T,\tau),\ldots,\Upsilon_5(s|\chi,T,\tau)\big)^{\mathsf{T}}\in\mathbb{R}^5$ *is the solution of the following auxiliary delay-differential system:*

$$\begin{aligned}\dot{\Upsilon}(s|\chi,T,\tau)=&\frac{\partial\tilde{h}(\tilde{x}(s),x(t_{delay}),T)}{\partial\tilde{x}}\Upsilon(s|\chi,T,\tau)+\frac{\partial\tilde{h}(\tilde{x}(s),x(t_{delay}),T)}{\partial x(t_{delay})}\Big[\Upsilon(s_{delay}|\chi,T,\tau)\\&+\frac{\partial x(t_{delay})}{\partial s_{delay}}\frac{\partial s_{delay}}{\partial T}\Big]+\frac{\partial\tilde{h}(\tilde{x}(s),x(t_{delay}),T)}{\partial T},s\in(0,1],\end{aligned} \tag{4.51}$$

with the initial condition

$$\Upsilon(s|\chi,T,\tau)=\mathbf{0}\in\mathbb{R}^5,\quad\forall s\le0. \tag{4.52}$$

Proof. From Theorem 4.2, we can conclude the desired conclusion immediately. □

Theorem 4.5. *For each pair* $(\chi,T)\in\mathscr{U}\times[a_3,b_3]$, $\frac{\partial\tilde{\Psi}(s|\chi,T,\tau)}{\partial u}=\Lambda(s|\chi,T,\tau)$, $s\in[0,1]$, *where* $\Lambda(s|\chi,T,\tau)=\big(\Lambda_{ij}(s|\chi,T,\tau)\big)_{5\times2}\in\mathbb{R}^{5\times2}$ *is the solution of the following auxiliary delay-differential system:*

$$\dot{\Lambda}(s|\chi,T,\tau)=\frac{\partial\tilde{h}(\tilde{x}(s),x(t_{delay}),T)}{\partial\tilde{x}}\Lambda(s|\chi,T,\tau)+\frac{\partial\tilde{h}(\tilde{x}(s),x(t_{delay}),T)}{\partial x(t_{delay})}\Lambda(s_{delay}|\chi,T,\tau),s\in(0,1], \tag{4.53}$$

with the initial condition

$$\Lambda(s|\chi,T,\tau)=\mathbf{0}\in\mathbb{R}^{5\times2},\quad\forall s\le0. \tag{4.54}$$

Proof. From Theorem 4.2, we can conclude the desired conclusion immediately. □

Remark 4.3. The computation of $\frac{\partial x(t_{delay})}{\partial s_{delay}}$ and $\frac{\partial s_{delay}}{\partial T}$ in Theorems 4.2 and 4.4 can be done as detailed in Remark 4.1 similarly.

Then, based on the above theorems, the following theorem gives the gradient formulas of $\widetilde{J}_1^{\beta}(\chi,T|\tau)$, $\widetilde{J}_2^{\beta}(\chi,T|\tau)$ and $\widetilde{C}_j^{\varepsilon,\gamma}(\chi,T|\tau)$, $j\in I_{10}$, with respect to χ and T.

Theorem 4.6. *Let* $(\chi,T)\in\mathscr{U}\times[a_3,b_3]$. *Then,*

$$\frac{\partial\widetilde{J}_1^{\beta}(\chi,T|\tau)}{\partial u}=-\frac{\Theta_3(1|\chi,T,\tau)}{T}+2\beta\frac{\tilde{\Psi}_3(1|\chi,T,\tau)}{T}\Lambda_3(1|\chi,T,\tau), \tag{4.55}$$

$$\frac{\partial \widetilde{J}_1^{\beta}(\chi,T|\tau)}{\partial T} = -\frac{\tilde{\psi}_3(1|\chi,T,\tau)}{T} + \frac{\tilde{x}_3(1|\chi,T,\tau)}{T^2} + 2\beta\frac{\tilde{\psi}_3(1|\chi,T,\tau)}{T}\Upsilon_3(1|\chi,T,\tau) - \frac{2\beta\tilde{\psi}_3^2(1|\chi,T,\tau)}{T^3}, \tag{4.56}$$

$$\frac{\partial \widetilde{J}_2^{\beta}(\chi,T|\tau)}{\partial u} = \frac{\partial u_2/\partial u - \Theta_2(1|\chi,T,\tau)}{T} + 2\beta\frac{\tilde{\psi}_3(1|\chi,T,\tau)}{T}\Lambda_3(1|\chi,T,\tau), \tag{4.57}$$

$$\frac{\partial \widetilde{J}_2^{\beta}(\chi,T|\tau)}{\partial T} = -\frac{\tilde{\psi}_2(1|\chi,T,\tau)}{T} - \frac{u_2 - \tilde{x}_2(1|\chi,T,\tau)}{T^2} + 2\beta\frac{\tilde{\psi}_2(1|\chi,T,\tau)}{T}\Upsilon_2(1|\chi,T,\tau) - \frac{2\beta\tilde{\psi}_2^2(1|\chi,T,\tau)}{T^3}, \tag{4.58}$$

$$\frac{\partial \widetilde{C}_j^{\varepsilon,\gamma}(\chi,T|\tau)}{\partial u} = \int_0^1 \frac{\varphi_{\varepsilon,j}(\widetilde{g}_j(\tilde{x}(s|\chi,T,\tau)))}{\partial \widetilde{g}_j}\frac{\partial \widetilde{g}_j}{\partial \tilde{x}}\Theta(1|\chi,T,\tau)\mathrm{d}s, \tag{4.59}$$

and

$$\frac{\partial \widetilde{C}_j^{\varepsilon,\gamma}(\chi,T|\tau)}{\partial T} = \int_0^1 \frac{\varphi_{\varepsilon,j}(\widetilde{g}_j(\tilde{x}(s|\chi,T,\tau)))}{\partial \widetilde{g}_j}\frac{\partial \widetilde{g}_j}{\partial \tilde{x}}\tilde{\psi}(1|\chi,T,\tau)\mathrm{d}s. \tag{4.60}$$

Proof. By the chain rule to (4.9) and Theorems 4.2-4.5, the desired conclusion can be obtained easily. □

We now propose the following algorithm for the computation of $\widetilde{J}_1^{\beta}(\chi,T|\tau)$ and its gradient $\frac{\partial \widetilde{J}_1^{\beta}(\chi,T|\tau)}{\partial T}$ corresponding to a candidate terminal time T.

Algorithm 4.1. **(i):** *Solve the delay-differential systems consisting of (4.8), (4.12), (4.13), (4.20), (4.21), (4.51), and (4.52) to obtain* $\tilde{x}_3(1|\chi,T,\tau)$, $\tilde{\psi}_3(1|\chi,T,\tau)$, $\Psi_3(1|\chi,T,\tau)$, *and* $\Upsilon_3(1|\chi,T,\tau)$ *by using any numerical integration scheme;*

(ii): *Use* $\tilde{x}_3(1|\chi,T,\tau)$ *and* $\tilde{\psi}_3(1|\chi,T,\tau)$ *to compute* $\widetilde{J}_1^{\beta}(\chi,T|\tau)$ *by using the formula (4.9);*

(iii): *Use* $\tilde{x}_3(1|\chi,T,\tau)$, $\tilde{\psi}_3(1|\chi,T,\tau)$, $\Psi_3(1|\chi,T,\tau)$, $\Upsilon_3(1|\chi,T,\tau)$, *and formula (4.56) to calculate* $\frac{\partial \widetilde{J}_1^{\beta}(\chi,T|\tau)}{\partial T}$.

Remark 4.4. The rest of the formulas in Theorem 4.6 can be similarly calculated.

4.4.3. *Parallel hybrid algorithm.* Genetic algorithms (hereafter, GA) are meta-heuristics inspired by the theory of evolution. They make use of the concepts, such as natural selection, reproduction, genetic heritage, and mutation, to search for "optimal solutions" of optimization problems; see, e.g., [51,52] and the references therein. In general, GA tends to perform well in exploration but badly in exploitation. For this, we will develop a hybrid GA (hereafter, HGA) algorithm to combine the advantage of the GA and some local search strategies. The GA is employed as the main frame of the proposed hybrid algorithm. A gradient-based algorithm based on Theorem 4.6 is employed as local search strategy.

To solve each of the resulting SOOPs, namely, $\widetilde{\mathbf{MOOP}}_{\mathbf{w}}^{\beta}$ and $\widetilde{\mathbf{MOOP}}_{\mathbf{mNBI}}^{\beta}$, we need to solve $\widetilde{\mathbf{MOOP}}_{(w_1,w_2)}^{\beta,\varepsilon,\gamma}$, $w_1,w_2 \geq 0, w_1+w_2=1$. For each of these problems, we need to numerically compute the value of the objective function given by (4.9) and the gradient formulas of the objective function and the constraint function. Thus, system (4.8) and the auxiliary dynamic systems in Theorems 4.2-4.5 are required to be solved numerically and repeatedly. If this process is to be carried out on a serial computer, it will be very time-consuming. Thus, a synchronous parallel hybrid algorithm is developed through a novel combination of GA and a gradient-based method with gradient formulas given in Theorem 4.6.

The parallel scheme used here is based on the Message Passing Interface (MPI) to provide a master-slave(Master-processor and Slave-processor) implementation. The Master-processor and Slave-processor scheme implemented here provides dynamic load balancing between the processors. In this implementation, one processor is always used as the Master-processor, and all remaining processors are used as Slave-processors. The Master-processor does computing and is also used to control the communication to and from the Slave-processors.

The parallel implementation starts with the Master-processor broadcasting some known parameters and experimental data to each Slave-processor. Each Slave-processor receives the information and generates one initial optimization vector. Each Slave-processor obtains the local best optimization vector in the current iteration and sends these information to Master-processor. The Master-processor obtains the global best optimization vector from all optimization vectors in the current iteration and broadcasts these information to each Slave-processor. Then, in the next iteration, each Slave-processor updates optimization vector based on the optimization vector of the preceding iteration. Repeat such iterations until our algorithm satisfies astopping criterion.

Details are given below:

Solve each of the resulting SOOPs, namely, $\widetilde{\mathbf{MOOP}}^{\beta,\varepsilon,\gamma}_{(w_1,w_2)}, w_1, w_2 \geq 0, w_1 + w_2 = 1$, by using gradient-based optimization and PGA techniques.

Algorithm 4.2. Step 1 : *Initialize data on Master-processor as follows:* $\beta, \varepsilon, \gamma, \mathscr{U}, a_3, b_3, W$, *and kinetic parameters in formulas (2.2)-(2.6) listed in Table 2.*

Step 2 : *Broadcast (***MPI_Bcast***) data on Master-processor to all Slave-processors.*

Step 3 : *Execute the following procedure on each Slave-processor, and denote the ID of processor as* $s \in I_{n_s}$.

- **Step 3.1** *: Set* $i \leftarrow 1$ *and randomly generate* n_w *genes with a uniform distribution on* $\mathscr{U} \times [a_3, b_3]$, *denoted by* $\sigma^i_{s,w}, \tilde{w} \in I_{n_w}$.
- **Step 3.2** *: Check and adjust. For each* $\tilde{w} \in I_{n_w}$,
  (1) based on formulas (4.19), check the values of $\widetilde{C}^{\varepsilon,\gamma}_j(\sigma^i_{s,\tilde{w}}|\tau), j \in I_{10}$.
  (2) if there exists $j \in I_{10}$ *such that* $\widetilde{C}^{\varepsilon,\gamma}_j(\sigma^i_{s,\tilde{w}}|\tau) < 0$ *(i.e.,* $\sigma^i_{s,\tilde{w}}$ *is infeasible for* $\widetilde{\mathbf{MOOP}}^{\beta,\varepsilon,\gamma}_{(w_1,w_2)}$, $w_1, w_2 \geq 0, w_1 + w_2 = 1$, *then, by using the gradient formulas given by (4.59)-(4.60) in Theorem 4.6, perform a gradient-ascent search (maximizing* $\widetilde{C}^{\varepsilon,\gamma}_j(\sigma^i_{s,\tilde{w}}|\tau)$*) to obtain a feasible point* $\breve{\sigma}^i_{s,\tilde{w}}$ *for* $\widetilde{\mathbf{MOOP}}^{\beta,\varepsilon,\gamma}_{(w_1,w_2)}, w_1, w_2 \geq 0, w_1 + w_2 = 1$*) (i.e.,* $\forall j \in I_{10}, \widetilde{C}^{\varepsilon,\gamma}_j(\breve{\sigma}^i_{s,\tilde{w}}|\tau) \geq 0$*), set* $\sigma^i_{s,\tilde{w}} \leftarrow \breve{\sigma}^i_{s,\tilde{w}}$ *and compute the corresponding objective function (4.14), denoted by* $\widetilde{J}^{\beta}_{(w_1,w_2)}(\sigma^i_{s,\tilde{w}}|\tau), w_1, w_2 \geq 0, w_1 + w_2 = 1$.
- **Step 3.3** *:**Crossover operator:***
  – *Set* $\tilde{l} \leftarrow 0$.
  – *For each* $\tilde{w}_1 \in I_{n_w}$, *if* $(r_{\tilde{w}_1} < p_c)$ *and (for certain* $\tilde{w}_2 \in I_{n_w}, \tilde{w}_2 \neq \tilde{w}_1, r_{\tilde{w}_2} < p_c$*), then*

$$\tilde{l} \leftarrow \tilde{l} + 1, \sigma^i_{s,n_w+\tilde{l}} \leftarrow \lambda \sigma^i_{s,\tilde{w}_1} + (1-\lambda)\sigma^i_{s,\tilde{w}_2},$$
$$\tilde{l} \leftarrow \tilde{l} + 1, \sigma^i_{s,n_w+\tilde{l}} \leftarrow (1-\lambda)\sigma^i_{s,\tilde{w}_1} + \lambda \sigma^i_{s,\tilde{w}_2}.$$

  It is easy to see that the generated offspring are still in $\mathscr{U} \times [a_3, b_3]$.
  – *Check and adjust. For each* $l \in I_{n_c}$,
    (1) based on formula (4.19), check the values of $\widetilde{C}^{\varepsilon,\gamma}_j(\sigma^i_{s,n_w+l}|\tau), j \in I_{10}$.
    (2) if there exits $j \in I_{10}$ *such that* $\widetilde{C}^{\varepsilon,\gamma}_j(\sigma^i_{s,n_w+l}|\tau) < 0$ *(i.e.,* σ^i_{s,n_w+l} *is infeasible for* $\widetilde{\mathbf{MOOP}}^{\beta,\varepsilon,\gamma}_{(w_1,w_2)}, w_1, w_2 \geq 0, w_1 + w_2 = 1$,*) then, by using the gradient formulas given by (4.59)-(4.60) in Theorem 4.6, perform a gradient-ascent search (maximizing* $\widetilde{C}^{\varepsilon,\gamma}_j(\sigma^i_{s,n_w+l}|\tau)$*) to obtain a feasible point* $\breve{\sigma}^i_{s,n_w+l}$ *for* $\widetilde{\mathbf{MOOP}}^{\beta,\varepsilon,\gamma}_{(w_1,w_2)}, w_1, w_2 \geq 0, w_1 + w_2 = 1$ *(i.e.,* $\forall j \in I_{10}, \widetilde{C}^{\varepsilon,\gamma}_j(\breve{\sigma}^i_{s,n_w+l}|\tau) \geq 0$*), set* $\sigma^i_{s,n_w+l} \leftarrow \breve{\sigma}^i_{s,n_w+l}$ *and compute the corresponding objective function (4.14), denoted by* $\widetilde{J}^{\beta}_{(w_1,w_2)}(\sigma^i_{s,n_w+l}|\tau), w_1, w_2 \geq 0, w_1 + w_2 = 1$.
- **Step 3.4** *: **Mutation operator:***

TABLE 3. Definitions of variables in Section 4.4.3.

Variable	Representation	Variable	Representation
σ	(χ, T)	ρ	the convergence tolerance of the objective function
P_{size}	the total number of genes in the population	$\lambda, r_{\tilde{w}_1}, r_{\tilde{w}_2}$	random numbers taken from $[0,1]$
n_s	the total number of processors	p_c	the crossover probability
$n_w \leftarrow P_{size}/n_s$	the number of genes on each processor	n_c	the total number of all offspring generated by crossover
i	the iteration index	p_m	the mutation probability
s	the processor index	r	random positive integer number of the set I_{n_c}
$\tilde{w}, \tilde{w}_1, \tilde{w}_2$	the gene index on each processor	n_m	the total number of mutated offspring
l	the iteration index in crossover operator	M	the maximum number of iterations
M_ρ	an integer for testing convergence	$\bar{\gamma}, \bar{\varepsilon}$	pre-specified positive constants

– *I_{n_c} is divided into two groups by* $\sigma^i_{s,n_w+*} \leftarrow \arg\min_{l\in n_c} \widetilde{J}^{\beta}_{(w_1,w_2)}(\sigma^i_{s,n_w+l}|\tau)$, *i.e.,*

$$I_{n_{c+}} \leftarrow \Big\{ l \in I_{n_c} \Big| |\sigma^i_{s,n_w+l} - \sigma^i_{s,n_w+*}| \geq \varepsilon \Big\},$$

$$I_{n_{c-}} \leftarrow \Big\{ l \in I_{n_c} \Big| |\sigma^i_{s,n_w+l} - \sigma^i_{s,n_w+*}| < \varepsilon \Big\},$$

where ε is a small positive number, and $|\cdot|$ represents the Euclidean norm.

– *Set* $\tilde{m} \leftarrow 0$.

– *For each* $l \in I_{n_c}$,

$$\begin{cases} \tilde{m} \leftarrow \tilde{m}+1, \sigma^i_{s,n_w+n_c+\tilde{m}} \leftarrow \tilde{\alpha}\sigma^i_{s,n_w+l} + (1-\tilde{\alpha})\sigma^i_{s,n_w+*}, \text{ if } r_l < p_m \text{ and } l \in I_{n_{c+}}, \\ \tilde{m} \leftarrow \tilde{m}+1, \sigma^i_{s,n_w+n_c+\tilde{m}} \leftarrow \tilde{\beta}\sigma^i_{s,n_w+l} + (1-\tilde{\beta})\sigma^i_{s,n_w+r}, \text{ if } r_l < p_m \text{ and } l \in I_{n_{c-}}, \end{cases}$$

where $\tilde{\alpha} \sim N(0,1)$, $\tilde{\beta} \sim N(0,1)$, and σ^i_{s,n_w+r} is generated randomly from $\mathscr{U} \times [a_3, b_3]$.

– *Check and adjust. For each* $m \in I_{n_m}$,

(1) based on formula (4.19), check the values of $\widetilde{C}^{\varepsilon,\gamma}_j(\sigma^i_{s,n_w+n_c+m}|\tau), j \in I_{10}$.

(2) if there exists $j \in I_{10}$ such that $\widetilde{C}^{\varepsilon,\gamma}_j(\sigma^i_{s,n_w+n_c+m}|\tau) < 0$ (i.e., σ^i_{s,n_w+n_c+m} is infeasible for $\widetilde{\mathbf{MOOP}}^{\beta,\varepsilon,\gamma}_{(w_1,w_2)}, w_1, w_2 \geq 0, w_1+w_2 = 1$, then, by using the gradient formulas given by (4.59)-(4.60) in Theorem 4.6, perform a gradient-ascent search (maximizing $\widetilde{C}^{\varepsilon,\gamma}_j(\sigma^i_{s,n_w+n_c+m}|\tau)$) to obtain a feasible point $\breve{\sigma}^i_{s,n_w+n_c+m}$ for $\widetilde{\mathbf{MOOP}}^{\beta,\varepsilon,\gamma}_{(w_1,w_2)}$, $w_1, w_2 \geq 0, w_1+w_2 = 1$ (i.e., $\forall j \in I_{10}, \widetilde{C}^{\varepsilon,\gamma}_j(\breve{\sigma}^i_{s,n_w+n_c+m}|\tau) \geq 0$), set $\sigma^i_{s,n_w+n_c+m} \leftarrow \breve{\sigma}^i_{s,n_w+n_c+m}$ and compute the corresponding objective function (4.14), denoted by

$$\widetilde{J}^{\beta}_{(w_1,w_2)}(\sigma^i_{s,n_w+n_c+m}|\tau), w_1, w_2 \geq 0, w_1+w_2 = 1.$$

- **Step 3.5** *: **Selection operator:***
  – *Based on the value of* $\widetilde{J}^{\beta}_{(w_1,w_2)}(\sigma^i_{s,d}|\tau)$, $d \in I_{n_w+n_c+n_m}$, *calculated in Steps 3.2∼3.4, select the best n_w genes from $n_w+n_c+n_m$ ones, denoted by $\sigma^{i+1}_{s,w}, w \in n_w$. The smaller the value of $\widetilde{J}^{\beta}_{(w_1,w_2)}(\sigma^i_{s,d}|\tau), d \in I_{n_w+n_c+n_m}$, the better the particles.*
  – *Calculate*

$$J^i_s \leftarrow \min_{d\in I_{n_w+n_c+n_m}} \widetilde{J}^{\beta}_{(w_1,w_2)}(\sigma^i_{s,d}|\tau), \ \sigma^i_s \leftarrow \arg\min_{d\in I_{n_w+n_c+n_m}} \widetilde{J}^{\beta}_{(w_1,w_2)}(\sigma^i_{s,d}|\tau).$$

- **Step 3.6** *: Starting with σ^i_s as the initial guess, the gradient formulas given by (4.55)–(4.58) in Theorem 4.6 to generate a descent direction.*
- **Step 3.7** *: Perform a line search along this direction to obtain the minimizer, still denoted by itself σ^i_s. After each iteration in this process, if $\widetilde{C}^{\varepsilon,\gamma}_j(\sigma^i_s|\tau) < 0, j \in I_{10}$, then perform a gradient-ascent search (maximizing $\widetilde{C}^{\varepsilon,\gamma}_j(\sigma^i_s|\tau)$) to obtain a feasible point $\breve{\sigma}^i_s$ for $\widetilde{\mathbf{MOOP}}^{\beta,\varepsilon,\gamma}_{(w_1,w_2)}, w_1, w_2 \geq 0$, $w_1+w_2 = 1$ (i.e., $\forall j \in I_{10}, \widetilde{C}^{\varepsilon,\gamma}_j(\breve{\sigma}^i_s|\tau) \geq 0$), set $\sigma^i_s \leftarrow \breve{\sigma}^i_s$. The minimal value of the corresponding objective function, is still denoted by itself J^i_s.*
- **Step 3.8** *: Send (**MPI_Send**) the values of J^i_s and σ^i_s from the sth Slave-processor to Master-processor.*

Step 4 : *Assign Master-processor to execute the tasks:*
  – *Receive (**MPI_Recv**) the values of J^i_s and σ^i_s from the sth Slave-processor.*
  – *Calculate*

$$J^{i+1} \leftarrow \min_{s\in I_{n_s}} J^i_s, \ \sigma^{i+1} \leftarrow \arg\min_{s\in I_{n_s}} J^i_s.$$

– *If* $\left(i \geq M\right)$ *or* $\left(i \geq M_\rho \text{ and } |J^{i+1} - J^i| < \rho\right)$, *then output* J^i *and* σ^i *stop. Otherwise, set* $i \leftarrow i+1$ *and go to Step 5.*

Step 5 : *Broadcast (***MPI_Bcast***)* σ^i *and* i *on Master-processor to all Slave-processors.*

Step 6 : *On each Slave-processor, use* σ^i *to replace the worst gene and go to Step 3.2.*

Remark 4.5. In Step 4, the algorithm stops when any of the following conditions holds:

- the maximum iteration M is reached, that is, $i \geq M$;
- the deviation between the best fitness of current iteration and that of the M_ρ previous iteration is less than ρ, that is, $i \geq M_\rho$ and $|J^{i+1} - J^i| < \rho$.

Combining Algorithm 4.2 with Remark 4.2, we propose the following algorithm for solving $\widetilde{\mathbf{MOOP}}^{\beta}_{(w_1,w_2)}$, $w_1, w_2 \geq 0, w_1 + w_2 = 1$.

Algorithm 4.3. Step 1 : *Choose initial values of* $\varepsilon > 0, \gamma > 0$.

Step 2 : *Solve* $\widetilde{\mathbf{MOOP}}^{\beta,\varepsilon,\gamma}_{(w_1,w_2)}, w_1, w_2 \geq 0, w_1 + w_2 = 1$, *using Algorithm 4.2 to obtain* $\sigma^*_{\varepsilon,\gamma}$.

Step 3 : *Check feasibility of* $\widetilde{g}_j(\tilde{x}(s|\sigma^*_{\varepsilon,\gamma}, \tau)) \leq 0$ *for* $s \in [0,1], j \in I_{10}$.

Step 4 : *If* $\sigma^*_{\varepsilon,\gamma}$ *is feasible, then go to Step 5. Otherwise, set* $\gamma \leftarrow \alpha_2\gamma$. *If* $\gamma > \bar{\gamma}$, *then go to Step 2. Otherwise, go to Step 6.*

Step 5 : *Set* $\varepsilon := \alpha_1\varepsilon$. *If* $\varepsilon > \bar{\varepsilon}$, *then go to Step 2. Otherwise, go to Step 6.*

Step 6 : *Output* $\sigma^*_{\varepsilon,\gamma}$ *and stop.*

Remark 4.6. Based on the obtained optimal solution of $\widetilde{\mathbf{MOOP}}^{\beta}_{(w_1,w_2)}, w_1, w_2 \geq 0, w_1 + w_2 = 1$, we can construct the form of Φ in (4.17) and solve $\widetilde{\mathbf{MOOP}}^{\beta}_{\mathbf{mNBI}}$ by the application of the exact penalty method [49], rather than the constraint transcription technique [50] and Theorem 4.6.

5. Numerical Results

The maximum productivity of 1,3-PD, the minimum consumption rate of glycerol, and the corresponding optimization in the $\mathbf{MOOP}^{\beta}$ are acquired by Algorithms 4.2, 4.3, and Theorem 4.6. The lower bounds and upper bounds for the initial state x_{01}, initial state x_{02}, and time durations together with the nominal time-delay, the initial terminal time and the history function are $a_1 = 0.02$, $b_1 = 400$, $a_2 = 0.202$, $b_2 = 500$, $a_3 = 5$, $b_3 = 10$, $\tau = 0.26h$, $T = 6$, and $\phi(t) = x_0 = (0.102gL^{-1}, 418.26087mmolL^{-1}, 0, 0, 0)^\intercal$, respectively [21]. The parameters selected in Algorithm 4.1, derived empirically after numerous experiments, are $P_{size} = 400$, $n_s = 80$, $M_\rho = 50$, $\rho = 0.01$, $p_c = 0.8$, $p_m = 0.2$, and $M = 600$. Specifically, $\varepsilon = 1 \times 10^{-2}$ and $\gamma = 1 \times 10^{-2}$ are taken as the initial values of the adjustable parameters, adjusting them at each step in terms of Algorithm 4.3. This algorithm involves reducing γ by a factor of 10 (i.e., $\alpha_2 = 10^{-1}$) or reducing both ε and γ by a factor of 10 (i.e., $\alpha_1 = \alpha_2 = 10^{-1}$) if constraints (4.10) and (4.11) do not hold. Algorithm 4.3 is terminated when $\varepsilon \leq \bar{\varepsilon} = 1 \times 10^{-7}$ and $\gamma \leq \bar{\gamma} = 1 \times 10^{-7}$.

System (4.8) together with auxiliary systems (4.12)-(4.13), (4.20)-(4.21), (4.49)-(4.50), (4.51)-(4.52), and (4.53)-(4.54) are solved by using the Euler integration scheme and the Lagrange interpolation. Lagrange interpolation [53] was utilized whenever auxiliary systems (4.12)-(4.13), (4.20)-(4.21), (4.49)-(4.50), (4.51)-(4.52) and (4.53)-(4.54) need the value of the state $x(t_{delay})$ or costate $\tilde{\Psi}(s_{delay}|\chi, T, \tau)$, $\Psi(s_{delay}|\chi, T, \tau)$, $\Theta(s_{delay}|\chi, T, \tau)$, $\Upsilon(s_{delay}|\chi, T, \tau)$, and $\Lambda(s_{delay}|\chi, T, \tau)$ at an intermediate time between two adjacent knot points.

$\mathbf{MOOP}^{\beta}$ is solved by the CWS as well as the mNBI approaches by using the single-objective optimization solver mentioned above, giving rise to 11 SOOPs. The Pareto sets depicted in Fig. 1 are obtained by these two methods. It is worth noting that a Pareto filter [28] is devised to eliminate all non-global Pareto solutions. This filter works by contrasting a point in the Pareto set with every other

generated points. A point will be removed if it is not a global Pareto solution. As it can be seen in Fig. 1, minimizing the negative productivity of $1,3$-PD and minimizing the consumption rate of substrate are conflicting objectives. Emphasizing on one of these two objectives can be achieved through a proper choice of the weights w_1 and $w_2 = 1 - w_1$. In particular, the minimization is focused on the negative productivity of $1,3$-PD when the value of w_1 approaches to one. On the other hand, minimizing is focused on the consumption rate of glycerol when the value of w_1 approaches to zero. Although both the approaches capture this feature, there are big differences in the precision of the resulting Pareto sets. Although an uniform spread is utilized to vary the weights in these two approaches, the results obtained by the mNBI approach clearly reveals a more uniform spread in the Pareto set than those obtained by the CWS approach. Consequently, the result obtained by the mNBI approach provides a more precise representation of Pareto set. Now, eleven points $A-K$ are chosen from the representation of the Pareto set obtained by the mNBI approach (see Fig. 1). The weights w_1, w_2, optimal solutions u^*, T^*, and optimal objectives $-G_1, G_2$ for points $A-K$ are listed in Table 4.

In our simulations, we perturb τ with various percentages as shown below

$$[-50\%, -45\%], [-45\%, -40\%], \ldots, [-5\%, 0], [0, 5\%], \ldots, [40\%, 45\%], [45\%, 50\%].$$

That is, τ is taken from the following disturbance intervals as follows:

$$[0.13, 0.143], [0.143, 0.156], \cdots, [0.247, 0.26], [0.26, 0.273], \ldots, [0.364, 0.377], [0.377, 0.39].$$

For each disturbance interval, we randomly generate 1000 of τ based on the procedure mentioned above and compute the corresponding value of the productivity of $1,3$-PD and the consumption rate of glycerol under the optimization strategy for the point F in Figs. 2-3. We observe that the change in the value of the productivity of $1,3$-PD and the consumption rate of glycerol with $\beta = 10^{-4}$ is much smaller than those obtained with $\beta = 0$. In other words, the value of the productivity of $1,3$-PD and the consumption rate of glycerol with $\beta = 10^{-4}$ is much more insensitivity with respect to the variation in τ than those obtained with $\beta = 0$.

TABLE 4. The weights w_1, w_2, optimal solutions u^*, T^* and optimal objectives $-G_1, G_2$ for points $A-K$ with mNBI.

	w_1	w_2	u^*	T^*	$-G_1$	G_2
A	0	1	(0.0200, 493.0167)$^\intercal$	9.1656	-8.6952	4.5809
B	0.1	0.9	(0.0200, 471.7586)$^\intercal$	7.6726	-19.5143	6.6317
C	0.2	0.8	(0.0200, 457.4018)$^\intercal$	6.3167	-30.8623	8.2618
D	0.3	0.7	(0.0200, 439.6159)$^\intercal$	5.8247	-41.5667	10.0841
E	0.4	0.6	(0.0406, 421.3284)$^\intercal$	5.3174	-52.5061	14.0345
F	0.5	0.5	(0.0679, 403.9863)$^\intercal$	5.0457	-62.0479	18.4152
G	0.6	0.4	(0.1082, 419.4859)$^\intercal$	6.0629	-70.5882	22.3258
H	0.7	0.3	(0.1745, 448.1653)$^\intercal$	6.8507	-77.8801	25.7295
I	0.8	0.2	(0.2020, 465.3412)$^\intercal$	7.4924	-83.2631	29.3383
J	0.9	0.1	(0.2020, 487.3359)$^\intercal$	7.9034	-86.5691	35.2486
K	1	0	(0.2020, 497.8235)$^\intercal$	8.1649	-88.6508	41.1432

6. DISCUSSIONS AND CONCLUSIONS

This paper considers a MOOP governed by nonlinear time-delay system arising in $1,3$-PD batch culture. The goal of the MOOP is to design an optimization scheme for balancing system cost and

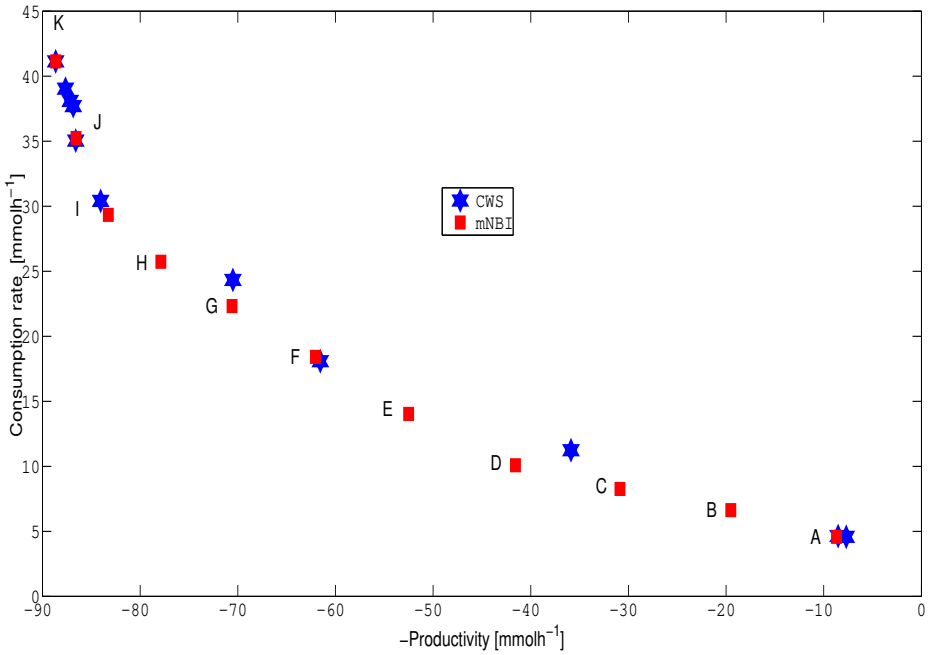

FIG. 1. Pareto sets generated by the mNBI and the CWS for $\beta = 10^{-4}$.

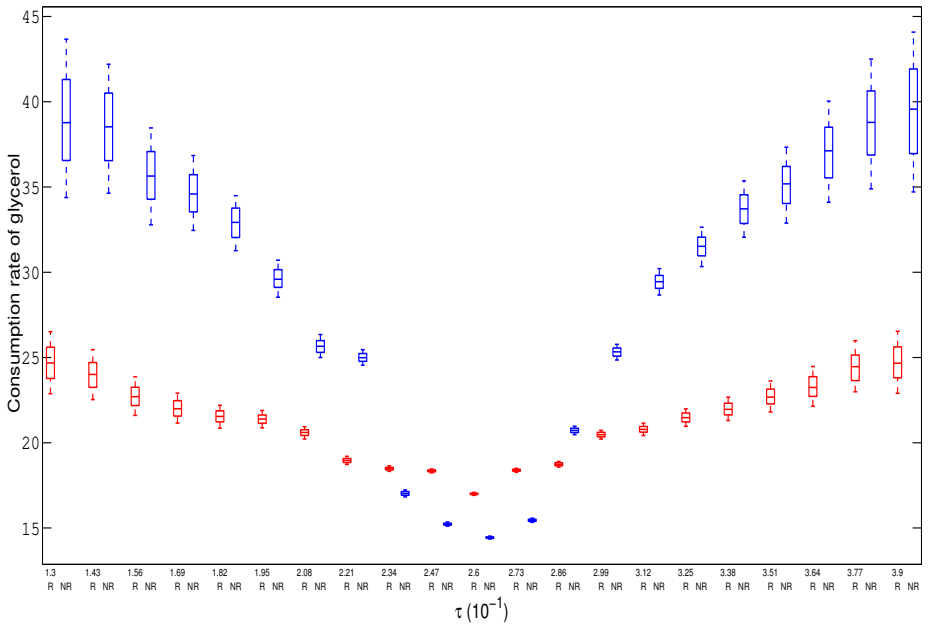

FIG. 2. Variation in the value of the consumption rate of glycerol due to the variation of τ for the point F. R: (red) $\beta = 10^{-4}$, NR: (blue) $\beta = 0$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

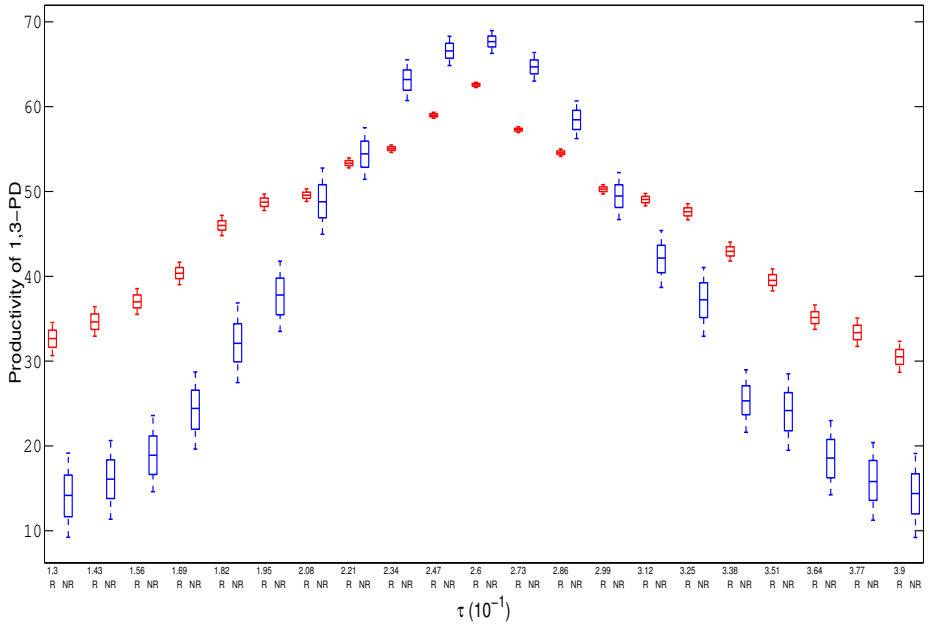

FIG. 3. Variation in the value of the productivity of 1,3-PD due to the variation of τ for the point F. R: (red) $\beta = 10^{-4}$, NR: (blue) $\beta = 0$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

system sensitivity. A parallel hybrid approach combining mNBI method and a single-objective solver is proposed for solving the MOOP. The effectiveness of the parallel hybrid approach is verified via numerical simulations. In closing, it should be noted that there may exist other uncertainties in the batch process, such as uncertainty in kinetic parameter. Thus, it is of great importance for the construction of optimal control schemes such that they are also insensitivity against these uncertainties. This is a future research topic.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (Grant Nos.: 11901075, 51979034, 11771008, 11701063, 71831002, 11871039, 61773086, and 12161076), the China Postdoctoral Science Foundation (Grant Nos.: 2019M661073 and 2019M651091), the Natural Science Foundation of Liaoning Province in China (Doctoral Start-up Foundation of Liaoning Province in China, Grant No.: 2020-BS-074), the Fundamental Scientific Research Projects of Higher Education Institutions of Liaoning Provincial Department of Education (Youth Project, LJKQZ2021011), the Fundamental Research Funds for the Central Universities, China (Grant Nos.: 3132022204, 3132021199, and DUT20YG125), the Australian Research Council (Grant No.: DP190103361), the Natural Science Foundation of Shandong Province in China (Grant Nos.: ZR2017MA005 and ZR2019MA031) and the Xinghai Project of Dalian Maritime University, China.

REFERENCES

[1] C. Groeger, W. Sabra, Z.P. 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.

[2] Z.P. Zeng, H. Biebl, Bulk-chemicals from biotechnology: the case of microbial production of 1,3-propanediol and the new trends. Adv. Biochem, Eng. Biotechnol. 74 (2002), 239-259.

[3] Z. Chen, F. Geng, A.P. Zeng, Protein design and engineering of a de novo pathway for microbial production of 1, 3-propanediol from glucose, Biotechnol. J. 10 (2015), 284-289.

[4] Y. An, B. Tan, L. Wang, L. Chang, Optimality condition and optimal control for a two-stage nonlinear dynamical system of microbial batch culture, Pacific J. Optim. 14 (2018), 1-13.

[5] Q. Yang, L. Wang, E.M. Feng, H.C. Yin, Z.L. Xiu, Identification and robustness analysis of Nonlinear Hybrid Dynamical system of genetic regulation in continuous culture, J. Indust. Manag. Optim. 16 (2020), 579-599.

[6] J.X. Ye, A. Li, J.G. Zhai, A measure of concentration robustness in a biochemical reaction network and its application on system identification, Appl. Math. Modell. 58 (2018), 270-280.

[7] Y.Q. Sun, J.T. Shen, L. Yan, J.J. Zhou, L.L. Jiang, Y. Chen, J.L. Yuan, E.M. Feng, Z.L. Xiu, Advances in bioconversion of glycerol to 1, 3-propanediol: prospects and challenges, Process Biochem. 71 (2018), 134-146.

[8] J. Yuan, C.Z. Wu, J.X. Ye, J. Xie, Robust identification of nonlinear state-dependent impulsive switched system with switching duration constraints, Nonlinear Anal. 36 (2020), 100879.

[9] J. Yuan, L. Wang, J. Xie, K.L. Teo, Optimal minimal variation control with quality constraint for Fed-Batch Fermentation processes involving multiple feeds. J. Franklin Inst. 357 (2020), 6571-6594.

[10] J. Wang, J.X. Ye, H.C. Yin, E.M. Feng, L. Wang, Sensitivity analysis and identification of kinetic parameters in batch fermentation of glycerol, J. Comput. Appl. Math. 236 (2012), 2268-2276.

[11] J. Zhang, J. Yuan, E. Feng, H. Yin, Z. Xiu, Strong stability of a nonlinear multi-stage dynamic system in batch culture of glycerol bioconversion to 1, 3-propanediol, Math. Model. Anal. 21 (2016), 159-173.

[12] J.L. Yuan, X. Zhang, X. Zhu, H.C. Yin, E.M. Feng, Z.L. Xiu, Modelling and pathway identification involving the transport mechanism of a complex metabolic system in batch culture, Commun. Nonlinear Sci. Numer. Simulat. 19 (2014), 2088-2103.

[13] X. Zhu, E.M. Feng, Joint estimation in batch culture by using unscented kalman filter, Biotechnol. Bioproc. Eng. 17 (2012), 1238-1243.

[14] L. Wang, E.M. Feng, Z.L. Xiu, Modeling nonlinear stochastic kinetic system and stochastic optimal control of microbial bioconversion process in batch culture, Nonlinear Anal. Model. 18 (2013), 99-111.

[15] J. Wang, J. Ye, E.M. Feng, H. Yin, Z. Xiu, Modeling and identification of a nonlinear hybrid dynamical system in batch fermentation of glycerol, Math. Comput. Model. 54 (2011), 618-624.

[16] J. Yuan, J. Xie, M. Huang, H.M. Fan, E.M. Feng, Z.L. Xiu, Robust optimal control problem with multiple characteristic time points in the objective for a batch nonlinear time-varying process using parallel global optimization, Optim. Eng. 21 (2020), 905-937.

[17] J. Yuan, L. Wang, J.G. Zhai, K.L. Teo, C.J. Yu, M. Huang, J. Xie, Robust optimal control for a batch nonlinear enzyme-catalytic switched time-delayed process with noisy output measurements, Nonlinear Anal. 41 (2021), 101059.

[18] C. Liu, Z. Gong, K.L. Teo, S. Wang, Modelling and optimal state-delay control in microbial batch process, Appl. Math. Model. 89 (2021), 792-801.

[19] D. Debeljković, Time-Delay Systems, InTech, Rijeka, 2011.

[20] C. Liu, M. Han, Z. Gong, K.L. Teo, Robust parameter estimation for constrained time-delay systems with inexact measurements, J. Indust. Manag. Optim. 17 (2021), 317-337.

[21] J. Yuan, C. Liu, X. Zhang, J. Xie, E. Feng, H. Yin, Z. Xiu, Optimal control of a batch fermentation process with nonlinear time-delay and free terminal time and cost sensitivity constraint, J. Process Contr. 44 (2016), 41-52.

[22] J.L. Yuan, L. Wang, X. Zhang, H.C. Yin, E.M. Feng, Z.L. Xiu, Parameter identification for a nonlinear enzyme-catalytic dynamic system with time-delays, J. Glob. Optim. 62 (2015), 791-810.

[23] V. Rehbock, K.L. Teo, L.S. Jennings, A computational procedure for suboptimal robust controls, Dynam. Control 2 (1992), 331-348.

[24] R.C. Loxton, K.L. Teo, V. Rehbock, Robust suboptimal control of nonlinear systems, Appl. Math. Comput. 217 (2011), 6566-6576.
[25] W. Wei, K.L. Teo, Z. Zhan, A numerical method for an optimal control problem with minimum sensitivity on coefficient variation, Appl. Math. Comput. 218 (2011), 1180-1190.
[26] Z. Gong, C. Liu, Y. Wang, Optimal control of switched systems with multiple time-delays and a cost on changing control, J. Ind. Manag. Optim. 14 (2018), 183-198.
[27] Z. Gong, C. Liu, K.L. Teo, S. Wang, Y. Wu, Numerical solution of free final time fractional optimal control problems, Appl. Math. Comput. 405 (2021), 126270.
[28] F. Logist, B. Houska, M. Diehl, J.F. Van Impe, Robust multi-objective optimal control of uncertain (bio) chemical processes, Chem. Eng. Sci. 66 (2011), 4670-4682.
[29] M. Vallerio, D. Telen, L. Cabianca, F. Manenti, J. Van Impe, F. Logist, Robust multi-objective dynamic optimization of chemical processes using the Sigma Point method, Chem. Eng. Sci. 140 (2016), 201-216.
[30] V. Bhaskar, S. Gupta, A. Ray, Applications of multi-objective optimization in chemical engineering, Rev. Chem. Eng. 16 (2000), 1-4.
[31] F. Logist, P.V. Erdeghem, J.V. Impe, Efficient deterministic multiple objective optimal control of (bio)chemical processes, Chem. Eng. Sci. 64 (2009), 2527-2538.
[32] F. Logist, B. Houska, M. Diehl, J.V. Impe, A toolkit for efficiently generating pareto sets in (bio)chemical multi-objective optimal control problems, Comput. Aided Chem. Eng. 28 (2010), 481-486.
[33] G. Xu, Y. Liu, Q. Gao, Multi-objective optimization of a continuous bio-dissimilation process of glycerol to 1, 3-propanediol, J. Biotechnol. 219 (2016), 59-71.
[34] F. Logist, B. Houska, M. Diehl, J.V. Impe, Fast Pareto set generation for nonlinear optimal control problems with multiple objectives, Struct. Multidiscip. Optim. 42 (2010), 591-603.
[35] C.Y. Kaya, H. Maurer, A numerical method for nonconvex multi-objective optimal control problems, Computat. Optim. Appl. 57 (2014), 685-702.
[36] L. Zadeh, Optimality and non-scalar-valued performance criteria, IEEE Trans. Autom. Control 8 (1963), 59-60.
[37] S.K. Shukla, On the normal boundary intersection method for generation of efficient front, In: International Conference on Computational Science 2007, pp. 310-317, Springer, Berlin Heidelberg, 2007.
[38] C.Y. Liu, Modelling and parameter identification for a nonlinear time-delay system in microbial batch fermentation, Appl. Math. Model. 37 (2013), 6899-6908.
[39] R.C. Loxton, K.L. Teo, V. Rehbock, F.F.C. Yiu, Optimal control problems with a continuous inequality constraint on the state and the control, Automatica 45 (2009), 2250-2257.
[40] Y. Gao, J. Lygeros, M. Quincampoix, On the reachability problem for uncertain hybrid systems, IEEE Trans. Auto. Control 52 (2007), 1572-1586.
[41] Y. Gao, J. Lygeros, M. Quincampoix, N. Seube, On the control of uncertain impulsive systems: Approximate stabilization and controlled invariance, Int. J. Control 77 (2004), 1393-1407.
[42] R. Loxton, Q. Lin, K.L. Teo, Switching time optimization for nonlinear switched systems: Direct optimization and the time-scaling transformation, Pac. J. Optim. 10 (2014), 537-560.
[43] C. Liu, R. Loxton, K.L. Teo, Dynamic optimization for switched time-delay systems with state-dependent switching conditions, SIAM J. Control Optim. 56 (2018), 3499-3523.
[44] Q. Lin, R. Loxton, K.L. Teo, The control parameterization method for nonlinear optimal control: A survey, J. Ind. Manag. Optim. 10 (2014), 275-309.
[45] C. Yu, Q. Lin, R. Loxton, K.L. Teo, G. Wang, A hybrid time-scaling transformation for time-delay optimal control problems, J. Optim. Theory Appl. 169 (2016), 876-901.
[46] I. Das, J.E. Dennis, Normal-boundary intersection: A new method for generating the Pareto surface in nonlinear multicriteria optimization problems, SIAM J. Optim. 8 (1998), 631-657.
[47] F.W. Gembicki, Performance and Sensitivity Optimization: A Vector Index Approach. PhD thesis, Case Western Reserve University, Cleveland, OH, 1974.
[48] A. Messac, A. Ismail-Yahaya, C.A. Mattson, The normalized normal constraint method for generating the Pareto frontier, Struct. Multidiscip. Optim. 25 (2003), 86-98.
[49] Q. Lin, R. Loxton, K.L. Teo, Y.H. Wu, C. Yu, A new exact penalty method for semi-infinite programming problems, J. Comput. Appl. Math. 261 (2014), 271-286.

[50] K.L. Teo, B. Li, C.J. Yu, V. Rehbock, Applied and Computational Optimal Control: A Control Parametrization Approach, Springer Optimization and Its Applications, 2021.
[51] R.K. Arakaki, F.L. Usberti, Hybrid genetic algorithm for the open capacitated arc routing problem, Comput. Oper. Res. 90 (2018), 221-231.
[52] M. Morini, S. Pellegrino, Personal income tax reforms: A genetic algorithm approach, Eur. J. Oper. Res. 264 (2018), 994-1004.
[53] J. Stoer, R. Bulirsch, Introduction to Numerical Analysis, Springer, New York, 1980.