Robust multi-objective optimal switching control arising in 1,3-propanediol microbial fed-batch process

Chongyang Liu[a,c], Zhaohua Gong[a], Kok Lay Teo[b,c], Jie Sun[c], Louis Caccetta[c]

[a]*School of Mathematics and Information Science, Shandong Institute of Business and Technology, Yantai 264005, Shandong, China*
[b]*School of Mathematics and Statistics, Changsha University of Science and Technology, Changsha 410114, China*
[c]*Department of Mathematics and Statistics, Curtin University, Perth 6845, Australia*

Abstract

This paper considers optimal control of glycerol producing 1,3-propanediol (1,3-PD) via microbial fed-batch fermentation. The fed-batch process is formulated as a nonlinear switched time-delay system. In general, the time-delay in the fed-batch process cannot be exactly estimated. Our goal is to design an optimal switching control scheme to simultaneously maximize 1,3-PD productivity and 1,3-PD yield under time-delay uncertainty. Accordingly, we propose a robust multi-objective optimal switching control model, in which two objectives, i.e., 1,3-PD productivity and 1,3-PD yield, and their sensitivities with respect to uncertain time-delay are considered in the objective vector. The control variables in this problem are the feeding rate of glycerol, the switching instants and the terminal time of the process. By introducing an auxiliary dynamic system to calculate the objective sensitivities and performing a time-scaling transformation, we obtain an equivalent multi-objective optimal switching control problem in standard form. We then convert the equivalent multi-objective optimal control problem into a sequence of single-objective optimal switching control problems by using a modified normal boundary intersection method. A novel gradient-based single-objective solver combining control parameterization with constraint transcription technique is developed to solve these resulting single-objective optimal control problems. Finally, numerical results are provided to verify the effectiveness of the proposed solution approach.

Key words: Switched time-delay system; Multi-objective optimal control; Normal boundary intersection; Control parameterization; Fed-batch process

1. Introduction

1,3-Propanediol (1,3-PD) is a promising bulk chemical which has attracted worldwide attention due to its enormous application in polymers, cosmetics,

foods, lubricants and medicines [1]. Currently, the market for 1,3-PD amounts to over 100 million pounds per year and is growing rapidly [2]. In general, production routes for 1,3-PD can be divided into two categories: chemical synthesis and microbial conversion. Compared with chemical synthesis, bioconversion of 1,3-PD via fermentation is particularly attractive in that the process is relatively easy and does not generate toxic byproducts. This conversion could also help to reduce glycerol surplus in the market [3]. Hence, improvement in the microbial production of 1,3-PD is of considerable importance to industries. To maximize the profit of 1,3-PD production, multi-objective optimization should be applied to the production process.

Glycerol can be converted to 1,3-PD by several microorganisms [4]. Among these, *Klebsiella pneumoniae* (*K. pneumoniae*) ferments glycerol to 1,3-PD in a good yield [5]. Glycerol fermentation to produce 1,3-PD by *K. pneumoniae* is a complex bioprocess [6]. Regarding the various fermentation techniques, fed-batch fermentation appears to be the most efficient cultivation method. The fed-batch fermentation is typically implemented by switching between batch mode and feeding mode. This switching manner can reduce effectively the substrate inhibition and improve the 1,3-PD productivity. Moreover, the concentration of substrate in a fed-batch process can be externally manipulated by using appropriate feeding rate profiles. As a result, optimal control of fed-batch processes has been a topic of research for many years [7, 8, 9].

The performance of optimal control depends on the accuracy of the process model. Recently, it is found that the fed-batch process can be accurately modelled as nonlinear dynamical systems. Under the assumption that the feed of glycerol only occurs at impulsive instants, the process is modelled as nonlinear impulsive system [10]. Taking the 1,3-PD concentration at the terminal time as the objective, optimal control of nonlinear impulsive systems is discussed in [11]. However, since the feeding rate of glycerol is finite, it is not reasonable to describe the actual fed-batch process as an impulsive dynamical system. In reality, the feed of glycerol is a continuous process. Thus, the fed-batch process is modelled as nonlinear switched systems [12, 13]. Optimal control of nonlinear switched systems is investigated in [14]. However, time-delays are ignored in the above nonlinear systems. In fact, like most real systems, fed-batch bioreactors are influenced by time-delays [15, 16]. Several reasons may be responsible for the occurrence of the delays in the fed-batch process: a cell has to undergo some change or growth process for which it needs some time before it reacts with others; the substrate and the products have to be transported across the cell membrane requiring a certain amount of time for transport; sometimes, either because of lack of knowledge or in order to reduce complexity it is appropriate to omit a number of intermediate steps in the reaction system for which the processing time is not negligible and has to be implemented as a delay [17, 18]. As a result, a nonlinear switched time-delay system is proposed in [19]. More recently, a published book [20] summarizes some optimal control results arising in 1,3-PD production processes. Although the results obtained are interesting, only one objective is considered in these optimal control problems and thus they all belong to single-objective optimal control (SOC) problems. Moreover, it is

2

difficult to determine the exact value of time-delay in the dynamic equation describing the fed-batch process and only nominal time-delay can be obtained using experimental data [19].

In this paper, we consider robust multi-objective optimal control (RMOC) of 1,3-PD fed-batch production in the presence of time-delay uncertainty. This fed-batch process is formulated as a nonlinear switched time-delay system. The optimal control problem is to design an optimal switching control scheme that maximizes 1,3-PD productivity and 1,3-PD yield, and also minimizes their sensitivities with respect to uncertain time-delay. Accordingly, we propose a robust multi-objective optimal switching control model that regards the feeding rate of glycerol, switching instants between batch and feeding modes, and the terminal time of the fermentation process as control variables and is subjected to continuous state inequality constraints. By the way, optimal control of switched systems is an important and challenging research topic for applied mathematicians [21, 22, 23, 24]. Nevertheless, optimal control for switched systems with time-delays are scarce in the literature. Necessary conditions for determining optimal switching times and/or optimal impulse magnitudes for such systems are derived in [25]. Switching time and parameter optimization for nonlinear switched systems with multiple time-delays is considered in [26]. However, no time-delay uncertainty is considered in these optimal switching control results. On the other hand, muti-objective optimal control (MOC) problems often arise in bioprocesses and have been extensively investigated; see, for example [27, 28, 29]. Multiple objective approaches are often employed to tackle these MOC problems: (i) scalarization methods, e.g., convex weighted sum (CWS) method [30], and normal boundary intersection method (NBI) [31], which transform MOC problem into a sequence of parametric SOC problems, and (ii) vectorization methods, e.g., genetic algorithm [32], and particle swarm optimization [33], which generate the Pareto set directly from the multi-objective formulation. It should be noted that scalarization approaches, compared with vectorization methods, can be combined with gradient-based deterministic optimization methods for finding optimal solutions to large-scale and highly constrained MOC problems in a fast and efficient way. As a result, scalarization approaches have been extensively used to solve MOC problems in biochemical processes [34, 35].

In this paper, by introducing an auxiliary time-delay system to calculate the objective sensitivities and performing a time-scaling transformation [36], we first transform the RMOC problem into an equivalent one in standard form. The equivalent problem is then converted into a sequence of parametric SOC problems by using a modified NBI method. The advantages of this method are that it can generate evenly distributed points in Pareto set and that it is weakly efficient for MOC problem. Incidentally, the existing single-objective solvers, including those developed in [28, 34, 35], only deal with SOC problems involving ordinary differential systems and thus cannot be used to solve the resulting SOC problems involving switched time-delay systems. For this reason, we approximate the resulting SOC by a sequence of parameter optimization problems through the application of the control parameterization method [37]. The continuous state inequality constraints are approximated as constraints in

3

canonical form by employing the constraint transcription technique [38]. The gradient formulas of the objectives and constraints with respect to the decision variables are also provided. On this basis, a novel gradient-based solver is developed to solve the resulting SOC problem. Finally, numerical results verify the effectiveness of the proposed solution approach.

The paper is organized as follows. Section 2 gives the process model describing the fed-batch process. Section 3 presents the RMOC problem. Section 4 gives the equivalent form of the RMOC problem. The numerical solution method for the equivalent problem is developed in Section 5. Numerical results are discussed in Section 6. Finally, Section 7 provides the main conclusions.

2. Process model

The fed-batch process for converting glycerol to 1,3-PD begins with a batch mode, then batch-fed glycerol and alkali are poured into the reactor in order to provide nutrition and maintain a suitable environment for the cells' growth. In particular, time-delays exist in the process [15, 16]. The process model of fed-batch process is based on the following assumptions:

Assumption 1. *The solution in the reactor is sufficiently well mixed so that the concentrations of reactants are uniform.*

Assumption 2. *The concentrations of biomass, substrate, 1,3-PD, acetate and ethanol in reactor at time t are determined by biomass concentration at time $t - \alpha$.*

Under the above Assumptions 1 and 2, the mass balance relationships for biomass, substrate and reaction products in the batch mode can be formulated as

$$\begin{cases} \dot{x}_1(t) = \mu(t)x_1(t - \alpha), \\ \dot{x}_2(t) = -q_2(t)x_1(t - \alpha), \\ \dot{x}_\ell(t) = q_\ell(t)x_1(t - \alpha), \quad \ell = 3, 4, 5, \\ \dot{x}_6(t) = 0, \end{cases}$$

where t denotes process time; $x_i(t)$, $i = 1, 2, 3, 4, 5, 6$, are, respectively, the extracellular concentrations of biomass, glycerol, 1,3-PD, acetic acid, ethanol, and the volume of culture fluid at time t in the fermentor; α is a time-delay; $\mu(t)$ is the specific growth rate of cells; $q_2(t)$ is the specific consumption rate of substrate; and $q_\ell(t)$, $\ell = 3, 4, 5$, are, respectively, the specific formation rates of the reaction products 1,3-PD, acetic acid and ethanol.

The mass balance relationships in the feeding mode can be expressed as

$$\begin{cases} \dot{x}_1(t) = \mu(t)x_1(t - \alpha) - D(t)x_1(t), \\ \dot{x}_2(t) = D(t)\Big(\dfrac{c_{s0}}{1 + r} - x_2(t)\Big) - q_2(t)x_1(t - \alpha), \\ \dot{x}_\ell(t) = q_\ell(t)x_1(t - \alpha) - D(t)x_\ell(t), \quad \ell = 3, 4, 5, \\ \dot{x}_6(t) = (1 + r)u(t), \end{cases}$$

4

where $c_{s0} = 10672 \text{mmolL}^{-1}$ is the concentration of the initial feed of glycerol; $r = 0.75$ is the velocity ratio of adding alkali to glycerol; $u(t)$ is the feeding rate of glycerol in the feeding mode; and $D(t)$ is the dilution rate at time t defined by

$$D(t) := \frac{(1+r)u(t)}{x_6(t)}.$$

Based on the work in [19], the specific growth rate of cells can be expressed as follows:

$$\mu(t) := \frac{\Delta_1 x_2(t)}{x_2(t) + k_1} \prod_{\ell=2}^{5} \left(1 - \frac{x_\ell(t)}{x_\ell^*}\right), \tag{1}$$

where Δ_1 is the maximum specific growth rate; k_1 is the Monod saturation constant; x_ℓ^* are the maximal residual concentrations of substrate and reaction products. The specific consumption rate of substrate is

$$q_2(t) := m_2 + \frac{\mu(t)}{Y_2} + \frac{\Delta_2 x_2(t)}{x_2(t) + k_2}, \tag{2}$$

where m_2 is the maintenance term of substrate consumption under substrate-limited conditions; Y_2 is the maximum growth yield; Δ_2 is the maximum increment of substrate consumption rate under substrate-sufficient conditions; and k_2 is the saturation constant for substrate. The specific formation rates of 1,3-PD and acetate are defined as

$$q_\ell(t) := m_\ell + Y_\ell \mu(t) + \frac{\Delta_\ell x_2(t)}{x_2(t) + k_\ell}, \quad \ell = 3, 4, \tag{3}$$

where m_3 and m_4 are the maintenance terms of product formations under substrate-limited conditions; Y_3 and Y_4 are the maximum product yields; Δ_3 and Δ_4 are the maximum increments of product formation rates under substrate-sufficient conditions; and k_3 and k_4 are saturation constants for 1,3-PD and acetate. The specific formation rate of ethanol can be described by

$$q_5(t) := q_2(t)\left(\frac{c_1}{c_2 + \mu(t)x_2(t)} + \frac{c_3}{c_4 + \mu(t)x_2(t)}\right), \tag{4}$$

where c_1, c_2, c_3 and c_4 are given parameters for the determination of the yield of ethanol on glycerol. Under anaerobic conditions at 37°C and pH 7.0, the maximal residual concentrations are $x_2^* = 2039 \text{mmolL}^{-1}$, $x_3^* = 1036 \text{mmolL}^{-1}$, $x_4^* = 1026 \text{mmolL}^{-1}$ and $x_5^* = 360.9 \text{mmolL}^{-1}$. The values of kinetic parameters in (1)-(4) are listed in Table 1.

Let N be the total number of batch and feeding modes. Then, the complete fed-batch process can be expressed as the following nonlinear switched time-delay system:

$$\begin{cases} \dot{x}(t) = f^i(x(t), x(t-\alpha), u(t)), & t \in (\tau_{i-1}, \tau_i], \quad i = 1, \dots, N, \\ x(t) = \phi(t), & t \leq \tau_0, \end{cases} \tag{5}$$

Table 1: The values of kinetic parameters in system (1)-(4) [19].

ℓ	m_ℓ	Y_ℓ	Δ_ℓ	k_ℓ	c_ℓ
1	-	-	0.8	0.28	0.025
2	1.927	0.0063	6.8489	17.7296	0.06
3	-3.2819	80.6096	10.3687	15.50	2.81
4	-0.97	33.07	5.74	85.71	65.5226

where $x(t) := (x_1(t), x_2(t), x_3(t), x_4(t), x_5(t), x_6(t))^\top$ is the state; $x(t-\alpha)$ is the delayed state; $u(t)$ is the control function;

$$f^{2j+1}(x(t), x(t-\alpha), u(t)) := \begin{pmatrix} \mu(t)x_1(t-\alpha) \\ -q_2(t)x_1(t-\alpha) \\ q_3(t)x_1(t-\alpha) \\ q_4(t)x_1(t-\alpha) \\ q_5(t)x_1(t-\alpha) \\ 0, \end{pmatrix}, \quad j = 0, \ldots, \left\lfloor \frac{N-1}{2} \right\rfloor;$$

$$f^{2j}(x(t), x(t-\alpha), u(t))$$
$$:= \begin{pmatrix} \mu(t)x_1(t-\alpha) - D(t)x_1(t) \\ D(t)\left(\frac{c_{s0}}{1+r} - x_2(t)\right) - q_2(t)x_1(t-\alpha) \\ q_3(t)x_1(t-\alpha) - D(t)x_3(t) \\ q_4(t)x_1(t-\alpha) - D(t)x_4(t) \\ q_5(t)x_1(t-\alpha) - D(t)x_5(t) \\ (1+r)u(t) \end{pmatrix}, \quad j = 1, \ldots, \left\lfloor \frac{N}{2} \right\rfloor;$$

$\lfloor \cdot \rfloor$ denotes the floor function; $\tau_1, \ldots, \tau_{N-1}$, are switching instants such that $0 = \tau_0 \le \tau_1 \le \cdots \le \tau_N = t_f$; t_f is the terminal time; and $\phi : R \to R^6$ is a given history function. Note that the number N is determined by the terminal time and the state of system (5) does not undergo jumps at the switching instants.

In system (5), the time-delay α may not be known exactly when the process is optimized, i.e., it can be uncertain. Nevertheless, the nominal time-delay α can be estimated using experimental data; see, for example [19]. Moreover, the concentrations of biomass, glycerol, reaction products and volume of culture fluid must be restricted to biological meaningful ranges. Hence, define

$$x(t) \in W := \prod_{\ell=1}^{6} [x_{*\ell}, x_\ell^*], \quad t \in [0, t_f], \tag{6}$$

where $x_{*1} = 0.01\mathrm{gL}^{-1}$, $x_{*2} = 15\mathrm{mmolL}^{-1}$, $x_{*3} = x_{*4} = x_{*5} = 0\mathrm{mmolL}^{-1}$ and $x_{*6} = 4\mathrm{L}$ are the lower thresholds for cell growth for biomass, glycerol, 1,3-PD, acetic acid, ethanol and volume of culture fluid, respectively; and $x_1^* = 6\mathrm{gL}^{-1}$, $x_6^* = 6.55\mathrm{L}$, x_2^*, x_3^*, x_4^* and x_5^* (as used in the formula for $\mu(t)$) are the corresponding upper thresholds.

3. Robust multi-objective optimal switching control problem

The control variables in the switched time-delay system (5) include the feeding rate of glycerol $u(t)$, switching instants τ_i, $i = 1, \ldots, N-1$, and the terminal time t_f. There are maximal and minimal time durations that are spent on each of the batch and feeding modes since biological considerations limit the rate of switching. On this basis, define

$$\Gamma := \{(\tau_1, \ldots, \tau_{N-1}, t_f)^\top \in R^N \mid \rho_i \leq \tau_i - \tau_{i-1} \leq \delta_i, \ i = 1, \ldots, N-1,$$
$$\rho_N \leq t_f \leq \delta_N\}, \tag{7}$$

where ρ_i and δ_i are lower and upper bounds, respectively. Any $\tau \in \Gamma$ is called an admissible vector of switching instants and terminal time.

During the batch mode, the feeding rate of glycerol is clearly zero. Moreover, during the feeding mode, the feeding rate of glycerol cannot be unbounded. Thus, define

$$U_i := \begin{cases} [a_i, b_i], & \text{if } i \text{ is even,} \\ \{0\}, & \text{if } i \text{ is odd,} \end{cases}$$

where $a_i > 0$ and $b_i > 0$ are lower and upper bounds for the feeding rate of glycerol during the ith feeding mode, respectively. Furthermore, we define

$$\mathcal{U} := \{u \in L_\infty([0, \delta_N], R) \mid u(t) \in U_i, \ i = 1, \ldots, N\}, \tag{8}$$

where $L_\infty([0, \delta_N], R)$ is the space of essentially bounded measurable functions from $[0, \delta_N]$ into R. Any $u \in \mathcal{U}$ is called an admissible control function. Accordingly, any pair $(u, \tau) \in \mathcal{U} \times \Gamma$ is called an admissible control pair. Let $x(\cdot|u, \tau, \alpha)$ denote the unique continuous solution of system (5) corresponding to the admissible control pair $(u, \tau) \in \mathcal{U} \times \Gamma$ and the nominal time-delay α.

Suppose that we are given a nominal time-delay α. The control objective in the fed-batch process is to maximize 1,3-PD productivity and 1,3-PD yield, simultaneously. The 1,3-PD productivity is the ratio between 1,3-PD mass and the process duration:

$$G_1(u, \tau|\alpha) := \frac{x_3(t_f|u, \tau, \alpha)x_6(t_f|u, \tau, \alpha)}{t_f}. \tag{9}$$

The 1,3-PD yield is the ratio of the amount of 1,3-PD formed and the amount of substrate consumed:

$$G_2(u, \tau|\alpha) := \frac{x_3(t_f|u, \tau, \alpha)x_6(t_f|u, \tau, \alpha)}{(x_6(t_f|u, \tau, \alpha) - \phi_6(0))c_{s0}}. \tag{10}$$

Let

$$J(u, \tau|\alpha) := (-G_1(u, \tau|\alpha), -G_2(u, \tau|\alpha))^\top \tag{11}$$

be the objective vector to be minimized. Then, taking into consideration of constraints (6), (7) and (8), we present the following MOC problem:

$$\text{(MOP)} \quad \min \quad J(u, \tau | \alpha)$$
$$\text{s.t.} \quad x(t | u, \tau, \alpha) \in W, \quad t \in [0, t_f],$$
$$(u, \tau) \in \mathcal{U} \times \Gamma.$$

In (MOP), the optimal control pair is determined under the assumption that the nominal time-delay α is exact. However, this is usually not the case in practice; the nominal time-delay is only an approximation of the true time-delay. Thus, we consider the following measure of productivity sensitivity with respect to the uncertain time-delay:

$$G_3(u, \tau | \alpha) := \left(\frac{\partial G_1(u, \tau | \alpha)}{\partial \alpha} \right)^2. \tag{12}$$

Clearly, (12) measures the rate at which productivity changes in response to small change in time-delay. Thus, a low value for productivity sensitivity indicates that the control pair is robust for productivity. Similarly, we present the following measure of yield sensitivity with respect to the uncertain time-delay:

$$G_4(u, \tau | \alpha) := \left(\frac{\partial G_2(u, \tau | \alpha)}{\partial \alpha} \right)^2. \tag{13}$$

Equation (13) measures the rate at which yield changes in response to small change in time-delay. Likewise, a low value for yield sensitivity implies that the control pair is robust for yield. We now propose the following modified objective vector that simultaneously minimizes (11), (12) and (13):

$$J^\beta(u, \tau | \alpha) := (-G_1(u, \tau | \alpha) + \beta G_3(u, \tau | \alpha), -G_2(u, \tau | \alpha) + \beta G_4(u, \tau | \alpha))^\top, \tag{14}$$

where $\beta \geq 0$ is a weight factor.

Thus, our RMOC problem can be stated as:

$$\text{(RMOP}^\beta) \quad \min \quad J^\beta(u, \tau | \alpha)$$
$$\text{s.t.} \quad x(t | u, \tau, \alpha) \in W, \quad t \in [0, t_f],$$
$$(u, \tau) \in \mathcal{U} \times \Gamma.$$

Note that (RMOP^β) becomes (MOP) when the weight factor β is equal to zero.

4. Problem transformation

Problem (RMOP^β) exhibits two non-standard aspects: (i) the objective vector contains non-standard sensitivity terms; and (ii) the terminal time is free instead of fixed. Thus, it is difficult to solve (RMOP^β) by numerical optimization algorithms [37, 39]. In the following subsections, we will circumvent these two challenges by problem transformation.

4.1. Computing sensitivity

Let

$$\zeta(t|u,\tau,\alpha) := \begin{cases} \dot{\phi}(t), & \text{if } t \leq \tau_0, \\ f^i(x(t), x(t-\alpha), u(t)), & \text{if } t \in (\tau_{i-1}, \tau_i], \ i \in \{1, \ldots, N\}. \end{cases}$$

Obviously, for almost all $t \in (-\infty, t_f]$, we have $\dot{x}(t|u,\tau,\alpha) = \zeta(t|u,\tau,\alpha)$.

For a given nominal time-delay α, we consider the following auxiliary time-delay system:

$$\begin{cases} \dot{\varphi}(t) = \dfrac{\partial f^i(x(t), x(t-\alpha), u(t))}{\partial x}\varphi(t) + \dfrac{\partial f^i(x(t), x(t-\alpha), u(t))}{\partial x(t-\alpha)} \\ \qquad \times \varphi(t-\alpha) - \dfrac{\partial f^i(x(t), x(t-\alpha), u(t))}{\partial x(t-\alpha)}\zeta(t-\alpha|u,\tau,\alpha), \\ \hspace{4cm} t \in (\tau_{i-1}, \tau_i], \quad i = 1, 2, \ldots, N, \\ \varphi(t) = 0, \quad t \leq \tau_0. \end{cases} \tag{15}$$

Let $\varphi(\cdot|u,\tau,\alpha)$ be the unique continuous solution of the auxiliary time-delay system (15) corresponding to the admissible control pair $(u,\tau) \in \mathcal{U} \times \Gamma$ and the nominal time-delay α. It can be shown (see, for example [40, 41]) that the state $x(t|u,\tau,\alpha)$ is differentiable with respect to α and

$$\frac{\partial x(t|u,\tau,\alpha)}{\partial \alpha} = \varphi(t|u,\tau,\alpha), \quad t \in (-\infty, t_f]. \tag{16}$$

On this basis, we obtain the following results which give the sensitivity terms with respect to α.

Theorem 1. *Let $(u,\tau) \in \mathcal{U} \times \Gamma$. Then, for a given nominal time-delay α,*

$$\left(\frac{\partial G_1(u,\tau|\alpha)}{\partial \alpha}\right)^2 = \frac{(\varphi_3(t_f)x_6(t_f) + \varphi_6(t_f)x_3(t_f))^2}{t_f^2},$$

and

$$\left(\frac{\partial G_2(u,\tau|\alpha)}{\partial \alpha}\right)^2 = \frac{(\varphi_3(t_f)x_6^2(t_f) - \phi_6(0)(\varphi_3(t_f)x_6(t_f) + \varphi_6(t_f)x_3(t_f)))^2}{(x_6(t_f) - \phi_6(0))^4 c_{s0}^2},$$

where $x(t_f) = x(t_f|u,\tau,\alpha)$; and $\varphi(t_f) = \varphi(t_f|u,\tau,\alpha)$.

Proof. Differentiate G_1 and G_2 with respect to α and then apply equation (16). $\square$

From Theorem 1, we see that the sensitivity terms of G_1 and G_2 with respect to α can be computed by solving the auxiliary time-delay system (15) with the switched time-delay system (5), simultaneously. Furthermore, based on Theorem 1, sensitivities (12) and (13) become

$$\bar{G}_3(u,\tau|\alpha) := \frac{(\varphi_3(t_f)x_6(t_f) + \varphi_6(t_f)x_3(t_f))^2}{t_f^2}, \tag{17}$$

9

and

$$\bar{G}_4(u,\tau|\alpha) := \frac{(\varphi_3(t_f)x_6^2(t_f) - \phi_6(0)(\varphi_3(t_f)x_6(t_f) + \varphi_6(t_f)x_3(t_f)))^2}{(x_6(t_f) - \phi_6(0))^4 c_{s0}^2}, \quad (18)$$

where $x(t_f) = x(t_f|u,\tau,\alpha)$; and $\varphi(t_f) = \varphi(t_f|u,\tau,\alpha)$.

4.2. Time-scaling transformation

The terminal time t_f in the switched time-delay system (5) and the auxiliary time-delay system (15) is free, that is, these systems must be integrated over a variable time horizon. For non-delay systems, this difficulty can be overcome by applying the following time-scaling transformation [36] to map interval $[0,t_f]$ into the fixed interval $[0,1]$:

$$t = t(s) = t_f s, \quad (19)$$

where $s \in [0,1]$ is a new time variable. Clearly, $s = 0$ corresponds to $t = 0$, $s = 1$ corresponds to $t = t_f$. This transformation is well known in the optimal control of non-delay systems. We now investigate its use for the switched time-delay system (5) and the auxiliary time-delay system (15). Let $\tilde{x}(s) := x(t_f s)$, $\tilde{u}(s) := u(t_f s)$ and $\theta_i := \tau_i/t_f$, $i = 1, \ldots, N$. Under the transformation (19), system (5) is transformed to the following switched time-delay system with fixed terminal time:

$$\begin{cases} \dot{\tilde{x}}(s) = \tilde{f}^i(\tilde{x}(s), \tilde{x}(s - t_f^{-1}\alpha), \tilde{u}(s), t_f), \quad s \in (\theta_{i-1}, \theta_i], \quad i = 1, \ldots, N, \\ \tilde{x}(s) = \tilde{\phi}(s), \quad s \leq \theta_0, \end{cases} \quad (20)$$

where $\theta_0 = 0$; $\tilde{f}^i(\tilde{x}(s), \tilde{x}(s - t_f^{-1}\alpha), \tilde{u}(s), t_f) := t_f f^i(\tilde{x}(s), \tilde{x}(s - t_f^{-1}\alpha), \tilde{u}(s))$; and $\tilde{\phi}(s) := \phi(t_f s)$. In addition, constraint (7) becomes

$$\Theta := \{\tilde{\theta} = (\theta_1, \ldots, \theta_{N-1}, t_f)^\top \in R^N| \; \tilde{\rho}_i \leq \theta_i - \theta_{i-1} \leq \tilde{\delta}_i, \; i = 1, \ldots, N-1,$$
$$\rho_N \leq t_f \leq \delta_N\},$$

where $\tilde{\rho}_i := \rho_i/t_f$; and $\tilde{\delta}_i := \delta_i/t_f$. For the set defined by (8), it becomes

$$\tilde{\mathcal{U}} := \{\tilde{u} \in L_\infty([0,1], R)| \; \tilde{u}(s) \in U_i, \; i = 1, \ldots, N\}. \quad (21)$$

Accordingly, let $\tilde{\varphi}(s) := \varphi(t_f s)$ and let $\partial \hat{x}$ denote differentiation with respect to $\tilde{x}(s - t_f^{-1}\alpha)$. Then, the auxiliary time-delay system (15) is transformed into an equivalent form given below:

$$\begin{cases} \dot{\tilde{\varphi}}(s) = \dfrac{\partial \tilde{f}^i(\tilde{x}(s), \tilde{x}(s - t_f^{-1}\alpha), \tilde{u}(s), t_f)}{\partial \tilde{x}} \tilde{\varphi}(s) \\ \qquad + \dfrac{\partial \tilde{f}^i(\tilde{x}(s), \tilde{x}(s - t_f^{-1}\alpha), \tilde{u}(s), t_f)}{\partial \hat{x}} \tilde{\varphi}(s - t_f^{-1}\alpha) \\ \qquad - t_f^{-1} \dfrac{\partial \tilde{f}^i(\tilde{x}(s), \tilde{x}(s - t_f^{-1}\alpha), \tilde{u}(s), t_f)}{\partial \hat{x}} \tilde{\zeta}(s - t_f^{-1}\alpha|\tilde{u}, \tilde{\theta}, \alpha), \\ \qquad\qquad\qquad\qquad s \in (\theta_{i-1}, \theta_i], \quad i = 1, \ldots, N, \\ \tilde{\varphi}(s) = 0, \quad s \leq \theta_0, \end{cases} \quad (22)$$

10

where

$$\tilde{\zeta}(s|\tilde{u}, \tilde{\theta}, \alpha)$$
$$:= \begin{cases} t_f \dot{\tilde{\phi}}(s), & \text{if } s \leq \theta_0, \\ \tilde{f}^i(\tilde{x}(s), \tilde{x}(s - t_f^{-1}\alpha), \tilde{u}(s), t_f), & \text{if } s \in (\theta_{i-1}, \theta_i], \ i \in \{1, \dots, N\}. \end{cases}$$

Denote the unique continuous solutions of systems (20) and (22) by $\tilde{x}(\cdot|\tilde{u}, \tilde{\theta}, \alpha)$ and $\tilde{\varphi}(\cdot|\tilde{u}, \tilde{\theta}, \alpha)$ corresponding to the admissible control pair $(\tilde{u}, \tilde{\theta}) \in \tilde{\mathcal{U}} \times \Theta$ and the nominal time-delay α on $(-\infty, 1]$. Then, constraint (6) changes to

$$\tilde{x}(s|\tilde{u}, \tilde{\theta}, \alpha) \in W, \quad s \in [0, 1]. \tag{23}$$

Moreover, substituting $\tilde{x}(\cdot|\tilde{u}, \tilde{\theta}, \alpha)$ and $\tilde{\varphi}(\cdot|\tilde{u}, \tilde{\theta}, \alpha)$ into equations (9), (10), (17) and (18) yields

$$\tilde{G}_1(\tilde{u}, \tilde{\theta}|\alpha) := \frac{\tilde{x}_3(1)\tilde{x}_6(1)}{t_f}, \tag{24}$$

$$\tilde{G}_2(\tilde{u}, \tilde{\theta}|\alpha) := \frac{\tilde{x}_3(1)\tilde{x}_6(1)}{(\tilde{x}_6(1) - \tilde{\phi}_6(0))c_{s0}}, \tag{25}$$

$$\tilde{G}_3(\tilde{u}, \tilde{\theta}|\alpha) := \frac{(\tilde{\varphi}_3(1)\tilde{x}_6(1) + \tilde{\varphi}_6(1)\tilde{x}_3(1))^2}{t_f^2}, \tag{26}$$

and

$$\tilde{G}_4(\tilde{u}, \tilde{\theta}|\alpha) := \frac{(\tilde{\varphi}_3(1)\tilde{x}_6^2(1) - \tilde{\phi}_6(0)(\tilde{\varphi}_3(1)\tilde{x}_6(1) + \tilde{\varphi}_6(1)\tilde{x}_3(1)))^2}{(\tilde{x}_6(1) - \tilde{\phi}_6(0))^4 c_{s0}^2}, \tag{27}$$

where $\tilde{x}(1) = \tilde{x}(1|\tilde{u}, \tilde{\theta}, \alpha)$; and $\tilde{\varphi}(1) := \tilde{\varphi}(1|\tilde{u}, \tilde{\theta}, \alpha)$. Then, objective vector (14) turns into

$$\tilde{J}^\beta(\tilde{u}, \tilde{\theta}|\alpha) := (-\tilde{G}_1(\tilde{u}, \tilde{\theta}|\alpha) + \beta\tilde{G}_3(\tilde{u}, \tilde{\theta}|\alpha), -\tilde{G}_2(\tilde{u}, \tilde{\theta}|\alpha) + \beta\tilde{G}_4(\tilde{u}, \tilde{\theta}|\alpha))^\top. \tag{28}$$

As a result, (RMOP^β) can be stated as the following equivalent problem:

$$(\text{ERMOP}^\beta) \quad \min \quad \tilde{J}^\beta(\tilde{u}, \tilde{\theta}|\alpha)$$
$$\text{s.t.} \quad \tilde{x}(s|\tilde{u}, \tilde{\theta}, \alpha) \in W, \quad s \in [0, 1],$$
$$(\tilde{u}, \tilde{\theta}) \in \tilde{\mathcal{U}} \times \Theta.$$

Obviously, (ERMOP^β) is a MOC problem with the terminal state cost and the fixed terminal time.

5. Numerical solution method

In this section, we will develop a numerical solution method combining a modified NBI with a single-objective optimal control solver to solve (ERMOP^β).

11

5.1. Modified normal boundary intersection

NBI method is one of the multi-objective optimization approaches developed in [42]. This approach is a powerful algorithm for producing an evenly distributed points in Pareto set and it is also proved that the method is independent of the relative scales of the objective functions [31].

Each objective function has a minimum value in its objective space called individual minima. This value is obtained independently by minimizing the following SOC problem:

$$
(\text{ERMOP}_k^\beta) \quad \min \quad \tilde{J}_k^\beta(\tilde{u}, \tilde{\theta}|\alpha), \quad k \in \{1, 2\} \tag{29}
$$
$$
\text{s.t.} \quad \tilde{x}(s|\tilde{u}, \tilde{\theta}, \alpha) \in W, \quad s \in [0, 1],
$$
$$
(\tilde{u}, \tilde{\theta}) \in \tilde{\mathcal{U}} \times \Theta,
$$

where $\tilde{J}_k^\beta$ is the kth component of the vector objective $\tilde{J}^\beta$. The obtained minima points for each objective function form important boundary points of the Pareto set, which is called the convex hull individual minima (CHIM). Utopia point is defined as the vector containing individual minima for each objective function (29):

$$
\tilde{J}^{\beta*} := (\tilde{J}_1^\beta(\tilde{u}^{1*}, \tilde{\theta}^{1*}|\alpha), \tilde{J}_2^\beta(\tilde{u}^{2*}, \tilde{\theta}^{2*}|\alpha))^\top,
$$

where $(\tilde{u}^{1*}, \tilde{\theta}^{1*})$ and $(\tilde{u}^{2*}, \tilde{\theta}^{2*})$ are the minimizers of the objectives $\tilde{J}_1^\beta$ and $\tilde{J}_2^\beta$, respectively. Then, the pay-off matrix Φ is defined as

$$
\Phi := \begin{bmatrix} 0 & \tilde{J}_1^\beta(\tilde{u}^{2*}, \tilde{\theta}^{2*}|\alpha) - \tilde{J}_1^\beta(\tilde{u}^{1*}, \tilde{\theta}^{1*}|\alpha) \\ \tilde{J}_2^\beta(\tilde{u}^{1*}, \tilde{\theta}^{1*}|\alpha) - \tilde{J}_2^\beta(\tilde{u}^{2*}, \tilde{\theta}^{2*}|\alpha) & 0 \end{bmatrix}. \tag{30}
$$

Let $\omega := (\omega_1, \omega_2)^\top$ be a vector such that $\omega_1 + \omega_2 = 1$ with $\omega_1, \omega_2 \geq 0$. $\Phi\omega$ describes a point in the CHIM. Moreover, let $e := (1, 1)^\top$ be the unit vector to CHIM. $-\Phi e$ defines the quasi-normal to the CHIM pointing towards the origin. The rationale behind the original NBI method is that the intersection point between the boundary of the objective space and the quasi-normal pointing towards the origin emanating from any point in the CHIM is expected to be Pareto optimal. Hereto, the multi-objective optimization problem is reformulated to maximize the distance from a point on the CHIM along the quasi-normal through this point, without violating the original constraints. Technically, this requirement of lying on the quasi-normal introduces additional equality constraints. Therefore, (ERMOP$^\beta$) can be reformulated as the following parametric SOC problem by the original NBI method:

$$
(\text{ERMOP}_{\text{oNBI}}^\beta) \quad \max \quad \nu
$$
$$
\text{s.t.} \quad \Phi\omega - \nu\Phi e = \tilde{J}^\beta(\tilde{u}, \tilde{\theta}|\alpha) - \tilde{J}^{\beta*},
$$
$$
\tilde{x}(s|\tilde{u}, \tilde{\theta}, \alpha) \in W, \quad s \in [0, 1],
$$
$$
(\tilde{u}, \tilde{\theta}) \in \tilde{\mathcal{U}} \times \Theta.
$$

The main problem with the original NBI method is that it may return non-Pareto optimal points. The method aims at getting boundary points rather

than Pareto-optimal points. Pareto-optimal points are a subset of boundary points and, therefore, these obtained points may or may not be Pareto-optimal points. Using the concept of the goal attainment approach [43], we present a modified NBI method in which the following parametric SOC problem is solved:

$$(\text{ERMOP}^{\beta}_{\text{mNBI}}) \qquad \max \quad \nu$$
$$\text{s.t.} \quad \Phi\omega - \nu\Phi e \geq \tilde{J}^{\beta}(\tilde{u}, \tilde{\theta}|\alpha) - \tilde{J}^{\beta*},$$
$$\tilde{x}(s|\tilde{u}, \tilde{\theta}, \alpha) \in W, \quad s \in [0, 1],$$
$$(\tilde{u}, \tilde{\theta}) \in \tilde{\mathcal{U}} \times \Theta.$$

This change in formulation increases the feasible space so as to check for points which dominate the points on the CHIM. By a similar argument as given for Lemma 1 in [44], it can be shown that if $(\nu^*, \tilde{u}^*, \tilde{\theta}^*)$ is an optimal solution of $(\text{ERMOP}^{\beta}_{\text{mNBI}})$ then $(\tilde{u}^*, \tilde{\theta}^*)$ is weakly efficient for (ERMOP^{β}). Moreover, a Pareto filter is designed to remove all non-global Pareto points whose definitions are given in [45]. This filter works by comparing a point in the Pareto set with every other generated point. If a point is not globally Pareto, then it is eliminated.

5.2. Single-objective optimal control solver

A sequence of resulting SOC problems must be solved when applying the modified NBI strategy. In this subsection, we shall develop a gradient-based single-objective optimal control solver combining the control parameterization method with a constraint transcription technique.

5.2.1. Control parameterization

The control parameterization method involves approximating the control by a linear combination of basis functions, thereby yielding an approximate optimization problem with a finite number of decision variables. Here, we focus on piecewise-constant basis functions, as these are the most widely used in practice. Other basis functions can be handled similarly; see, for example [37].

For each $p_i \geq 1$, $i \in \{1, \ldots, N\}$, let the time subinterval $[\theta_{i-1}, \theta_i]$ be partitioned into n_{p_i} subintervals with $n_{p_i} + 1$ partition points denoted by

$$\theta_0^{p_i}, \ldots, \theta_{n_{p_i}}^{p_i}, \ \theta_0^{p_i} = \theta_{i-1}, \ \theta_{n_{p_i}}^{p_i} = \theta_i, \ \text{and} \ \theta_{l-1}^{p_i} < \theta_l^{p_i}.$$

Let n_{p_i} be chosen such that $n_{p_i+1} \geq n_{p_i}$. With piecewise-constant basis functions, the control $\tilde{u}$ is approximated as follows:

$$\tilde{u}(s) \approx \tilde{u}^p(s) = \sigma^{p_i, l}, \ s \in (\theta_{l-1}^{p_i}, \theta_l^{p_i}], \ l = 1, \ldots, n_{p_i}, \ i = 1, \ldots, N,$$

where $\sigma^{p_i, l}$ is the value of the control on the lth subinterval of $(\theta_{i-1}, \theta_i]$ which is to be chosen optimally. The approximate piecewise-constant control can be written as follows:

$$\tilde{u}^p(s|\sigma^p, \tilde{\theta}) := \sum_{i=1}^{N} \sum_{l=1}^{n_{p_i}} \sigma^{p_i, l} \chi_{(\theta_{l-1}^{p_i}, \theta_l^{p_i}]}(s), \quad s \in [0, 1], \tag{31}$$

where

$$\sigma^p := ((\sigma^{p_1})^\top, \ldots, (\sigma^{p_N})^\top)^\top$$

with $\sigma^{p_i} := (\sigma^{p_i,1}, \ldots, \sigma^{p_i,n_{p_i}})^\top$; and for interval I, $\chi_I : R \to R$ is the characteristic function defined by

$$\chi_I(s) := \begin{cases} 1, & \text{if } s \in I, \\ 0, & \text{if } s \notin I. \end{cases}$$

From (21), we obtain the following bound constraints on σ^p:

$$\sigma^{p_i,l} \in U_i, \quad l = 1, \ldots, n_{p_i}, \ i = 1, \ldots, N. \tag{32}$$

Let Ξ^p denote the set of all σ^p satisfying constraint (32).

Substituting (31) into system (20) yields

$$\begin{cases} \dot{\tilde{x}}(s) = \tilde{f}^i(\tilde{x}(s), \tilde{x}(s - t_f^{-1}\alpha), \sigma^{p_i,l}, t_f), & s \in (\theta_{l-1}^{p_i}, \theta_l^{p_i}], \\ \qquad\qquad\qquad\qquad\qquad l = 1, \ldots, n_{p_i}, \ i = 1, \ldots, N, \\ \tilde{x}(s) = \tilde{\phi}(s), \quad s \le \theta_0, \end{cases} \tag{33}$$

Furthermore, substituting (31) into the auxiliary time-delay system (22) gives

$$\begin{cases} \dot{\tilde{\varphi}}(s) = \dfrac{\partial \tilde{f}^i(\tilde{x}(s), \tilde{x}(s - t_f^{-1}\alpha), \sigma^{p_i,l}, t_f)}{\partial \tilde{x}} \tilde{\varphi}(s) \\ \qquad + \dfrac{\partial \tilde{f}^i(\tilde{x}(s), \tilde{x}(s - t_f^{-1}\alpha), \sigma^{p_i,l}, t_f)}{\partial \hat{x}} \tilde{\varphi}(s - t_f^{-1}\alpha) \\ \qquad - t_f^{-1} \dfrac{\partial \tilde{f}^i(\tilde{x}(s), \tilde{x}(s - t_f^{-1}\alpha), \sigma^{p_i,l}, t_f)}{\partial \hat{x}} \tilde{\zeta}^p(s - t_f^{-1}\alpha | \sigma^p, \tilde{\theta}, \alpha), \\ \qquad\qquad s \in (\theta_{l-1}^{p_i}, \theta_l^{p_i}], \quad l = 1, \ldots, n_{p_i}, \quad i = 1, 2, \ldots, N, \\ \tilde{\varphi}(s) = 0, \quad s \le \theta_0, \end{cases} \tag{34}$$

where

$$\begin{aligned} & \tilde{\zeta}^p(s|\sigma^p, \tilde{\theta}, \alpha) \\ & := \begin{cases} t_f \dot{\tilde{\phi}}(s), & \text{if } s \le \theta_0, \\ \tilde{f}^i(\tilde{x}(s), \tilde{x}(s - t_f^{-1}\alpha), \sigma^p, t_f), & \text{if } s \in (\theta_{i-1}, \theta_i], \ i \in \{1, \ldots, N\}. \end{cases} \end{aligned} \tag{35}$$

Let $\tilde{x}^p(\cdot | \sigma^p, \tilde{\theta}, \alpha)$ and $\tilde{\varphi}^p(\cdot | \sigma^p, \tilde{\theta}, \alpha)$ be the unique continuous solutions of systems (33) and (34) corresponding to the parameterized control pair $(\sigma^p, \tilde{\theta}) \in \Xi^p \times \Theta$ and the nominal time-delay α. Then, constraint (23) becomes

$$\tilde{x}^p(s|\sigma^p, \tilde{\theta}, \alpha) \in W, \quad s \in [0, 1]. \tag{36}$$

The objective vector (28) changes to

$$\begin{aligned} & \tilde{J}^{\beta,p}(\sigma^p, \tilde{\theta}|\alpha) \\ & := (-\tilde{G}_1^p(\sigma^p, \tilde{\theta}|\alpha) + \beta \tilde{G}_3^p(\sigma^p, \tilde{\theta}|\alpha), -\tilde{G}_2^p(\sigma^p, \tilde{\theta}|\alpha) + \beta \tilde{G}_4^p(\sigma^p, \tilde{\theta}|\alpha))^\top, \end{aligned} \tag{37}$$

14

where $\tilde{G}_1^p$, $\tilde{G}_2^p$, $\tilde{G}_3^p$ and $\tilde{G}_4^p$ are obtained by substituting the solutions $\tilde{x}^p$ and $\tilde{\varphi}^p$ into equations (24), (25), (26) and (27), respectively.

Thus, by the control parameterization technique, (ERMOP$_{\mathrm{mNBI}}^{\beta}$) can be approximated by the following parameter optimization problem:

$$(\mathrm{ERMOP}_{\mathrm{mNBI}}^{\beta,p}) \qquad \max \quad \nu \tag{38}$$

$$\text{s.t.} \quad \Phi^p \omega - \nu \Phi^p e \geq \tilde{J}^{\beta,p}(\sigma^p, \tilde{\theta}|\alpha) - \tilde{J}^{\beta,p*}, \tag{39}$$

$$\tilde{x}^p(s|\sigma^p, \tilde{\theta}, \alpha) \in W, \quad s \in [0,1],$$

$$(\sigma^p, \tilde{\theta}) \in \Xi^p \times \Theta,$$

where $(\sigma^{p1*}, \tilde{\theta}^{1*})$ and $(\sigma^{p2*}, \tilde{\theta}^{2*})$ are the minimizers of the objectives $\tilde{J}_1^{\beta,p}$ and $\tilde{J}_2^{\beta,p}$, respectively; $\tilde{J}^{\beta,p*} := (\tilde{J}_1^{\beta,p}(\sigma^{p1*}, \tilde{\theta}^{1*}|\alpha), \tilde{J}_2^{\beta,p}(\sigma^{p2*}, \tilde{\theta}^{2*}|\alpha))^{\top}$; and Φ^p is obtained by replacing $(\tilde{u}^{1*}, \tilde{u}^{2*})$ in (30) with $(\sigma^{p1*}, \sigma^{p2*})$.

5.2.2. Constraint transcription

In essence, constraint (36) is a continuous state inequality constraint. This type of constraint often arises in many real-world applications [46, 47]. Notwithstanding, it is difficult to deal with such constraints in the numerical computation. To achieve this, an integral penalty method [48], an equivalent end-point constraints method [49] and an exact penalty method [50] were introduced. However, a common characteristic of all these techniques is that the penalty terms or end-point constraints introduced have zero gradients with respect to optimization variables at the optimal solution. This, in turn, can result in a reduced convergence rate near the optimal solution. We will apply a constraint transcription technique [38] to approximate constraint (36) by a sequence of constraints in canonical form.

Let

$$h_\ell(\tilde{x}^p(s|\sigma^p, \tilde{\theta}, \alpha)) := x_\ell^* - \tilde{x}_\ell^p(s|\sigma^p, \tilde{\theta}, \alpha),$$

$$h_{6+\ell}(\tilde{x}^p(s|\sigma^p, \tilde{\theta}, \alpha)) := \tilde{x}_\ell(s|\sigma^p, \tilde{\theta}, \alpha) - x_{*\ell}, \quad \ell = 1, \ldots, 6.$$

Then, constraint (36) can be transformed into the following canonical equality constraint:

$$\sum_{\imath=1}^{12} \int_0^1 \min\{0, h_\imath(\tilde{x}^p(s|\sigma^p, \tilde{\theta}, \alpha))\} ds = 0. \tag{40}$$

However, (40) is a non-smooth constraint since $\min\{0, \cdot\}$ is non-differentiable at the origin. Thus, gradient-based optimization methods, such as sequential quadratic programming (SQP) [51], will typically struggle to handle this constraint. As a result, we consider the following smooth approximation of the function $\min\{0, \cdot\}$:

$$\min\{0, \eta\} \approx \pi_\epsilon(\eta) := \begin{cases} \eta, & \text{if } \eta < -\epsilon, \\ -(\eta - \epsilon)^2/4\epsilon, & \text{if } -\epsilon \leq \eta \leq \epsilon, \\ 0, & \text{otherwise}, \end{cases}$$

15

where $\epsilon > 0$ is an adjustable parameter. It is easy to verify that π_ϵ is continuously differentiable and non-positive. Substituting $\min\{0, \eta\} \approx \pi_\epsilon(\eta)$ into the left-hand side of (40) gives the following constraint:

$$\sum_{i=1}^{12} \int_0^1 \pi_\epsilon(h_i(\tilde{x}^p(s|\sigma^p, \tilde{\theta}, \alpha)))ds = 0. \tag{41}$$

However, this constraint is not suitable for numerical computation since it does not satisfy the linear independence constraint qualification—a fundamental condition required for the convergence of nonlinear programming algorithms. Therefore, we further approximate (41) by

$$H^{\epsilon,\gamma}(\sigma^p, \tilde{\theta}|\alpha) := \gamma + \sum_{i=1}^{12} \int_0^1 \pi_\epsilon(h_i(\tilde{x}^p(s|\sigma^p, \tilde{\theta}, \alpha)))ds \geq 0, \tag{42}$$

where $\gamma > 0$ is an adjustable parameter. Replacing (36) in $(\text{ERMOP}^p_{\text{mNBI}})$ with (42) yields standard nonlinear programming problems. In addition, it can be shown, as in [38], that for each $\epsilon > 0$, there exists a corresponding $\gamma(\epsilon) > 0$ such that whenever $0 < \gamma < \gamma(\epsilon)$, constraint (42) implies constraint (36). From the results in [38], we see 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 fraction of time horizon, then it appears that $\gamma = O(\epsilon)$ should be chosen. On the other hand, if the constraint is active only over a very small fraction of time horizon, then $\gamma = O(\epsilon^2)$.

5.2.3. Gradient formulas

To solve each of the resulting SOC problems using the control parameterization in conjunction with the constraint transcription technique in the previous Subsections 5.2.1 and 5.2.2, we need to solve a sequence of standard nonlinear programming problems using optimization methods. Among various optimization techniques, a gradient-based optimization, such as SQP, is much more efficient [38]. For this, we need the gradients of the objective vector (37) and constraint function (42) with respect to σ^p and $\tilde{\theta}$. These gradients are presented in the following theorems whose proofs are similar to those given for Theorems 1-3 in [36]. Note that, in the sequel, we will use $\partial\bar{x}$ to denote differentiation with respect to $\tilde{x}(s + t_f^{-1}\alpha)$

First, define

$$\hat{f}^{i,l}(s|\sigma^p, \tilde{\theta}, \alpha) := \tilde{f}^i(\tilde{x}^p(s|\sigma^p, \tilde{\theta}, \alpha), \tilde{x}^p(s - t_f^{-1}\alpha|\sigma^p, \tilde{\theta}, \alpha), \sigma^{p_i,l}, t_f), \tag{43}$$

$$\psi(s|\sigma^p, \tilde{\theta}, \alpha)$$
$$:= \begin{cases} 0, & \text{if } s \leq \theta_0, \\ \bar{f}^{i,l}(s|\sigma^p, \tilde{\theta}, \alpha), & \text{if } s \in (\theta_{l-1}^{p_i}, \theta_l^{p_i}], \ l \in \{1, \ldots, n_{p_i}\}, \ i \in \{1, \ldots, N\}, \end{cases} \tag{44}$$

16

and

$$
\begin{aligned}
&\varpi(s|\sigma^p, \tilde{\theta}, \alpha) \\
&:= \begin{cases} t_f \dfrac{d^2\phi(s)}{ds^2}, & \text{if } s \le \theta_0, \\ \vartheta(s|\sigma^p, \tilde{\theta}, \alpha) & \text{if } s \in (\theta_{l-1}^{p_i}, \theta_l^{p_i}], \ l \in \{1, \dots, n_{p_i}\}, \ i \in \{1, \dots, N\}, \end{cases}
\end{aligned} \tag{45}
$$

where

$$
\begin{aligned}
&\bar{f}^{i,l}(s|\sigma^p, \tilde{\theta}, \alpha) \\
&= \frac{\partial \hat{f}^{i,l}(s|\sigma^p, \tilde{\theta}, \alpha)}{\partial \tilde{x}} \tilde{\varphi}^p(s|\sigma^p, \tilde{\theta}, \alpha) + \frac{\partial \hat{f}^{i,l}(s|\sigma^p, \tilde{\theta}, \alpha)}{\partial \hat{x}} \tilde{\varphi}^p(s - t_f^{-1}\alpha|\sigma^p, \tilde{\theta}, \alpha) \\
&\quad - t_f^{-1} \frac{\partial \hat{f}^{i,l}(s|\sigma^p, \tilde{\theta}, \alpha))}{\partial \hat{x}} \tilde{\zeta}^p(s - t_f^{-1}\alpha|\sigma^p, \tilde{\theta}, \alpha);
\end{aligned} \tag{46}
$$

and

$$
\begin{aligned}
&\vartheta(s|\sigma^p, \tilde{\theta}, \alpha) \\
&= \frac{\partial \hat{f}^{i,l}(s|\sigma^p, \tilde{\theta}, \alpha)}{\partial \tilde{x}} \hat{f}^{i,l}(s|\sigma^p, \tilde{\theta}, \alpha) + \frac{\partial \hat{f}^{i,l}(s|\sigma^p, \tilde{\theta}, \alpha)}{\partial \hat{x}} \tilde{\zeta}^p(s - t_f^{-1}\alpha|\sigma^p, \tilde{\theta}, \alpha).
\end{aligned}
$$

Obviously, for almost all $s \in (-\infty, 1]$, we have $\dot{\tilde{\varphi}}^p(s|\sigma^p, \tilde{\theta}, \alpha) = \psi(s|\sigma^p, \tilde{\theta}, \alpha)$ and $\dot{\tilde{\zeta}}^p(s|\sigma^p, \tilde{\theta}, \alpha) = \varpi(s|\sigma^p, \tilde{\theta}, \alpha)$. Furthermore, define

$$
\mathcal{S}_l^{p_i}(\theta, t_f, \alpha) := (\theta_{l-1}^{p_i} - t_f^{-1}\alpha, \theta_l^{p_i} - t_f^{-1}\alpha].
$$

Then, the following theorem gives the gradient formulas of $\tilde{J}_1^{\beta,p}(\sigma^p, \tilde{\theta}|\alpha)$ with respect to σ^p and $\tilde{\theta}$.

Theorem 2. *Let* $(\sigma^p, \tilde{\theta}) \in \Xi^p \times \Theta$. *Then, for a given nominal time-delay* α,

$$
\begin{aligned}
\frac{\partial \tilde{J}_1^{\beta,p}(\sigma^p, \tilde{\theta}|\alpha)}{\partial \tilde{\theta}_i} &= \mu^\top(\tilde{\theta}_i) \bar{f}^{i,n_{p_i}}(\tilde{\theta}_i|\sigma^p, \tilde{\theta}, \alpha) - \mu^\top(\tilde{\theta}_i) \bar{f}^{i+1,1}(\tilde{\theta}_i|\sigma^p, \tilde{\theta}, \alpha) \\
&\quad + \lambda^\top(\tilde{\theta}_i) \hat{f}^{i,n_{p_i}}(\tilde{\theta}_i|\sigma^p, \tilde{\theta}, \alpha) - \lambda^\top(\tilde{\theta}_i) \hat{f}^{i+1,1}(\tilde{\theta}_i|\sigma^p, \tilde{\theta}, \alpha), \\
&\hspace{6cm} i = 1, \dots, N-1, \\
\frac{\partial \tilde{J}_1^{\beta,p}(\sigma^p, \tilde{\theta}|\alpha)}{\partial \sigma^{p_i,l}} &= \int_{\theta_{l-1}^{p_i}}^{\theta_l^{p_i}} \left\{ \mu^\top(s) \frac{\partial \bar{f}^{i,l}(s|\sigma^p, \tilde{\theta}, \alpha)}{\partial \sigma^{p_i,l}} + \lambda^\top(s) \frac{\partial \hat{f}^{i,l}(s|\sigma^p, \tilde{\theta}, \alpha)}{\partial \sigma^{p_i,l}} \right\} ds, \\
&\hspace{5cm} l = 1, \dots, n_{p_i}, \ i = 1, \dots N,
\end{aligned}
$$

17

and

$$\frac{\partial \tilde{J}_1^{\beta,p}(\sigma^p, \tilde{\theta}|\alpha)}{\partial t_f} = t_f^{-2} \big[\tilde{x}_3^p(1)\tilde{x}_6^p(1) - 2\beta\tilde{\varphi}_3^p(1)\tilde{x}_6^p(1) - 2\beta\tilde{\varphi}_6^p(1)\tilde{x}_3^p(1) \big]$$

$$+ \sum_{i=1}^{N} \sum_{l=1}^{n_{p_i}} \int_{\theta_{l-1}^{p_i}}^{\theta_l^{p_i}} \left\{ \mu^\top(s)\mathcal{L}_1(s) + \lambda^\top(s)\mathcal{L}_2(s) \right\} ds$$

$$+ \sum_{i=1}^{N} \sum_{l=1}^{n_{p_i}} \int_{\theta_{l-1}^{p_i} - t_f^{-1}\alpha}^{\theta_l^{p_i} - t_f^{-1}\alpha} \left\{ \left\{ \mu^\top(s + t_f^{-1}\alpha)\mathcal{L}_3(s) + \lambda^\top(s + t_f^{-1}\alpha)\mathcal{L}_4(s) \right\} \right.$$

$$\left. \times s\dot{\phi}(t_f s)\chi_{(-\infty,0]}(s) \right\} ds,$$

where $\tilde{x}^p(1) := \tilde{x}^p(1|\sigma^p, \tilde{\theta}, \alpha)$; $\tilde{\varphi}^p(\cdot) := \tilde{\varphi}^p(\cdot|\sigma^p, \tilde{\theta}, \alpha)$; $\hat{f}^{i,l}(\cdot|\sigma^p, \tilde{\theta}, \alpha)$ and $\bar{f}^{i,l}(\cdot|\sigma^p, \tilde{\theta}, \alpha)$ are as defined in (43) and (46), respectively; $\mathcal{L}_1(\cdot)$, $\mathcal{L}_2(\cdot)$, $\mathcal{L}_3(\cdot)$ and $\mathcal{L}_4(\cdot)$ are, respectively, as defined in (A.1), (A.2), (A.3) and (A.4) in Appendix; and $\mu(\cdot) := \mu(\cdot|\sigma^p, \tilde{\theta}, \alpha)$ and $\lambda(\cdot) := \lambda(\cdot|\sigma^p, \tilde{\theta}, \alpha)$ are the solutions of the following costate system:

$$\dot{\mu}(s) = -\left(\sum_{i=1}^{N} \sum_{l=1}^{n_{p_i}} \frac{\partial \hat{f}^{i,l}(s|\sigma^p, \tilde{\theta}, \alpha)}{\partial \tilde{x}} \right)^\top \mu(s) \chi_{\mathcal{S}_l^{p_i}(\theta, t_f, 0)}(s)$$

$$- \left(\sum_{i=1}^{N} \sum_{l=1}^{n_{p_i}} \frac{\partial \hat{f}^{i,l}(s + t_f^{-1}\alpha|\sigma^p, \tilde{\theta}, \alpha)}{\partial \tilde{x}} \right)^\top \mu(s + t_f^{-1}\alpha) \chi_{\mathcal{S}_l^{p_i}(\theta, t_f, \alpha)}(s),$$

$$s \in [0, 1],$$

$$\dot{\lambda}(s) = -\sum_{i=1}^{N} \sum_{l=1}^{n_{p_i}} \left(\left(\frac{\partial \hat{f}^{i,l}(s|\sigma^p, \tilde{\theta}, \alpha)}{\partial \tilde{x}} \right)^\top \lambda(s) + \left(\frac{\partial^2 \hat{f}^{i,l}(s|\sigma^p, \tilde{\theta}, \alpha)}{\partial \tilde{x}^2} \tilde{\varphi}^p(s) \right. \right.$$

$$+ \frac{\partial^2 \hat{f}^{i,l}(s|\sigma^p, \tilde{\theta}, \alpha)}{\partial \tilde{x}\partial \hat{x}} \tilde{\varphi}^p(s - t_f^{-1}\alpha) - t_f^{-1} \frac{\partial}{\partial \tilde{x}} \left\{ \frac{\partial \hat{f}^{i,l}(s|\sigma^p, \tilde{\theta}, \alpha)}{\partial \hat{x}} \right.$$

$$\left. \left. \times \tilde{\zeta}^p(s - t_f^{-1}\alpha|\sigma^p, \tilde{\theta}, \alpha) \right\} \right)^\top \mu(s) \right) \chi_{\mathcal{S}_l^{p_i}(\theta, t_f, 0)}(s)$$

$$- \sum_{i=1}^{N} \sum_{l=1}^{n_{p_i}} \left(\left(\frac{\partial \hat{f}^{i,l}(s + t_f^{-1}\alpha|\sigma^p, \tilde{\theta}, \alpha)}{\partial \tilde{x}} \right)^\top \lambda(s + t_f^{-1}\alpha) \right.$$

$$+ \left(\frac{\partial^2 \hat{f}^{i,l}(s + t_f^{-1}\alpha|\sigma^p, \tilde{\theta}, \alpha)}{\partial \tilde{x}^2} \tilde{\varphi}^p(s) + \frac{\partial^2 \hat{f}^{i,l}(s + t_f^{-1}\alpha|\sigma^p, \tilde{\theta}, \alpha)}{\partial \tilde{x}\partial \bar{x}} \tilde{\varphi}^p(s + t_f^{-1}\alpha) \right.$$

$$\left. \left. - t_f^{-1} \frac{\partial}{\partial \tilde{x}} \left\{ \frac{\partial \hat{f}^{i,l}(s + t_f^{-1}\alpha|\sigma^p, \tilde{\theta}, \alpha)}{\partial \tilde{x}} \tilde{\zeta}^p(s|\sigma^p, \tilde{\theta}, \alpha) \right\} \right)^\top \mu(s + t_f^{-1}\alpha) \right)$$

$$\times \chi_{\mathcal{S}_l^{p_i}(\theta, t_f, \alpha)}(s), \qquad\qquad s \in [0, 1],$$

with the terminal conditions

$$\mu(1) = (0, 0, \partial \tilde{J}_1^{\beta,p}(\sigma^p, \tilde{\theta}|\alpha)/\partial \tilde{\varphi}_3, 0, 0, \partial \tilde{J}_1^{\beta,p}(\sigma^p, \tilde{\theta}|\alpha)/\partial \tilde{\varphi}_6)^\top, \qquad (47)$$

$$\lambda(1) = (0, 0, \partial \tilde{J}_1^{\beta,p}(\sigma^p, \tilde{\theta}|\alpha)/\partial \tilde{x}_3, 0, 0, \partial \tilde{J}_1^{\beta,p}(\sigma^p, \tilde{\theta}|\alpha)/\partial \tilde{x}_6)^\top, \qquad (48)$$

$$\mu(s) = \lambda(s) = (0, 0, 0, 0, 0, 0)^\top, \quad s > 1.$$

The next theorem gives the gradient formulas of $\tilde{J}_2^p(\sigma^p, \tilde{\theta}|\alpha)$ with respect to σ^p and $\tilde{\theta}$.

Theorem 3. *Let* $(\sigma^p, \tilde{\theta}) \in \Xi^p \times \Theta$. *Then, for a given nominal time-delay* α,

$$\frac{\partial \tilde{J}_2^{\beta,p}(\sigma^p, \tilde{\theta}|\alpha)}{\partial \tilde{\theta}_i} = \bar{\mu}^\top(\tilde{\theta}_i) \bar{f}^{i,n_{p_i}}(\tilde{\theta}_i|\sigma^p, \tilde{\theta}, \alpha) - \bar{\mu}^\top(\tilde{\theta}_i) \bar{f}^{i+1,1}(\tilde{\theta}_i|\sigma^p, \tilde{\theta}, \alpha)$$

$$+ \bar{\lambda}^\top(\tilde{\theta}_i) \hat{f}^{i,n_{p_i}}(\tilde{\theta}_i|\sigma^p, \tilde{\theta}, \alpha) - \bar{\lambda}^\top(\tilde{\theta}_i) \hat{f}^{i+1,1}(\tilde{\theta}_i|\sigma^p, \tilde{\theta}, \alpha),$$

$$i = 1, \ldots, N - 1,$$

$$\frac{\partial \tilde{J}_2^{\beta,p}(\sigma^p, \tilde{\theta}|\alpha)}{\partial \sigma^{p_i,l}} = \int_{\theta_{l-1}^{p_i}}^{\theta_l^{p_i}} \left\{ \bar{\mu}^\top(s) \frac{\partial \bar{f}^{i,l}(s|\sigma^p, \tilde{\theta}, \alpha)}{\partial \sigma^{p_i,l}} + \bar{\lambda}^\top(s) \frac{\partial \hat{f}^{i,l}(s|\sigma^p, \tilde{\theta}, \alpha)}{\partial \sigma^{p_i,l}} \right\} ds,$$

$$l = 1, \ldots, n_{p_i}, \ i = 1, \ldots N,$$

and

$$\frac{\partial \tilde{J}_2^{\beta,p}(\sigma^p, \tilde{\theta}|\alpha)}{\partial t_f} = \sum_{i=1}^N \sum_{l=1}^{n_{p_i}} \int_{\theta_{l-1}^{p_i}}^{\theta_l^{p_i}} \left\{ \bar{\mu}^\top(s) \mathcal{L}_1(s) + \bar{\lambda}^\top(s) \mathcal{L}_2(s) \right\} ds$$

$$+ \sum_{i=1}^N \sum_{l=1}^{n_{p_i}} \int_{\theta_{l-1}^{p_i} - t_f^{-1}\alpha}^{\theta_l^{p_i} - t_f^{-1}\alpha} \left\{ \left\{ \bar{\mu}^\top(s + t_f^{-1}\alpha) \mathcal{L}_3(s) \right. \right.$$

$$+ \bar{\lambda}^\top(s + t_f^{-1}\alpha) \mathcal{L}_4(s) \right\} s\dot{\phi}(t_f s) \chi_{(-\infty,0]}(s) \Big\} ds,$$

where $\hat{f}^{i,l}(\cdot|\sigma^p, \tilde{\theta}, \alpha)$ *and* $\bar{f}^{i,l}(\cdot|\sigma^p, \tilde{\theta}, \alpha)$ *are as defined in (43) and (46), respectively;* $\mathcal{L}_1(\cdot)$, $\mathcal{L}_2(\cdot)$, $\mathcal{L}_3(\cdot)$ *and* $\mathcal{L}_4(\cdot)$ *are, respectively, as defined in (A.1), (A.2), (A.3) and (A.4) in Appendix; and* $\bar{\mu}(\cdot) := \bar{\mu}(\cdot|\sigma^p, \tilde{\theta}, \alpha)$ *and* $\bar{\lambda}(\cdot) := \bar{\lambda}(\cdot|\sigma^p, \tilde{\theta}, \alpha)$ *are the solutions of the following costate system:*

$$\dot{\bar{\mu}}(s) = -\left(\sum_{i=1}^N \sum_{l=1}^{n_{p_i}} \frac{\partial \hat{f}^{i,l}(s|\sigma^p, \tilde{\theta}, \alpha)}{\partial \tilde{x}} \right)^\top \bar{\mu}(s) \chi_{\mathcal{S}_l^{p_i}(\theta, t_f, 0)}(s)$$

$$- \left(\sum_{i=1}^N \sum_{l=1}^{n_{p_i}} \frac{\partial \hat{f}^{i,l}(s + t_f^{-1}\alpha|\sigma^p, \tilde{\theta}, \alpha)}{\partial \tilde{x}} \right)^\top \bar{\mu}(s + t_f^{-1}\alpha) \chi_{\mathcal{S}_l^{p_i}(\theta, t_f, \alpha)}(s),$$

$$s \in [0, 1],$$

$$
\begin{aligned}
\dot{\bar{\lambda}}(s) = {} & -\sum_{i=1}^{N}\sum_{l=1}^{n_{p_i}}\left(\left(\frac{\partial \hat{f}^{i,l}(s|\sigma^p,\tilde{\theta},\alpha)}{\partial \tilde{x}}\right)^{\top}\bar{\lambda}(s) + \left(\frac{\partial^2 \hat{f}^{i,l}(s|\sigma^p,\tilde{\theta},\alpha)}{\partial \tilde{x}^2}\tilde{\varphi}^p(s)\right.\right. \\
& + \frac{\partial^2 \hat{f}^{i,l}(s|\sigma^p,\tilde{\theta},\alpha)}{\partial \tilde{x}\partial \hat{x}}\tilde{\varphi}^p(s - t_f^{-1}\alpha) - t_f^{-1}\frac{\partial}{\partial \tilde{x}}\left\{\frac{\partial \hat{f}^{i,l}(s|\sigma^p,\tilde{\theta},\alpha)}{\partial \hat{x}}\right. \\
& \left.\left.\times \tilde{\zeta}^p(s - t_f^{-1}\alpha|\sigma^p,\tilde{\theta},\alpha)\right\}\right)^{\top}\bar{\mu}(s)\Bigg)\chi_{\mathcal{S}_l^{p_i}(\theta,t_f,0)}(s) \\
& -\sum_{i=1}^{N}\sum_{l=1}^{n_{p_i}}\left(\left(\frac{\partial \hat{f}^{i,l}(s + t_f^{-1}\alpha|\sigma^p,\tilde{\theta},\alpha)}{\partial \tilde{x}}\right)^{\top}\bar{\lambda}(s + t_f^{-1}\alpha)\right. \\
& + \left(\frac{\partial^2 \hat{f}^{i,l}(s + t_f^{-1}\alpha|\sigma^p,\tilde{\theta},\alpha)}{\partial \tilde{x}^2}\tilde{\varphi}^p(s) + \frac{\partial^2 \hat{f}^{i,l}(s + t_f^{-1}\alpha|\sigma^p,\tilde{\theta},\alpha)}{\partial \tilde{x}\partial \bar{x}}\right. \\
& \left.\times \tilde{\varphi}^p(s + t_f^{-1}\alpha) - t_f^{-1}\frac{\partial}{\partial \tilde{x}}\left\{\frac{\partial \hat{f}^{i,l}(s + t_f^{-1}\alpha|\sigma^p,\tilde{\theta},\alpha)}{\partial \tilde{x}}\tilde{\zeta}^p(s|\sigma^p,\tilde{\theta},\alpha))\right\}\right)^{\top} \\
& \left.\times \bar{\mu}(s + t_f^{-1}\alpha)\right)\chi_{\mathcal{S}_l^{p_i}(\theta,t_f,\alpha)}(s), \qquad\qquad s \in [0,1],
\end{aligned}
$$

with the terminal conditions

$$
\bar{\mu}(1) = (0,0,\partial \tilde{J}_2^{\beta,p}(\sigma^p,\tilde{\theta}|\alpha)/\partial \tilde{\varphi}_3, 0,0, \partial \tilde{J}_2^{\beta,p}(\sigma^p,\tilde{\theta}|\alpha)/\partial \tilde{\varphi}_6)^{\top}, \qquad (49)
$$

$$
\bar{\lambda}(1) = (0,0,\partial \tilde{J}_2^{\beta,p}(\sigma^p,\tilde{\theta}|\alpha)/\partial \tilde{x}_3, 0,0, \partial \tilde{J}_2^{\beta,p}(\sigma^p,\tilde{\theta}|\alpha)/\partial \tilde{x}_6)^{\top}, \qquad (50)
$$

$$
\bar{\mu}(s) = \bar{\lambda}(s) = (0,0,0,0,0,0)^{\top}, \quad s > 1.
$$

Note that the explicit formulas $\partial \tilde{J}_1^{\beta,p}/\partial \tilde{\varphi}_3$, $\partial \tilde{J}_1^{\beta,p}/\partial \tilde{\varphi}_6$, $\partial \tilde{J}_1^{\beta,p}/\partial \tilde{x}_3$, $\partial \tilde{J}_1^{\beta,p}/\partial \tilde{x}_6$, $\partial \tilde{J}_2^{\beta,p}/\partial \tilde{\varphi}_3$, $\partial \tilde{J}_2^{\beta,p}/\partial \tilde{\varphi}_6$, $\partial \tilde{J}_2^{\beta,p}/\partial \tilde{x}_3$ and $\partial \tilde{J}_2^{\beta,p}/\partial \tilde{x}_6$ in (47)-(50) are as given in (A.5)-(A.12) in Appendix, respectively. The last theorem gives the gradient formulas of the constraint function $H^{\epsilon,\gamma}(\sigma^p,\tilde{\theta}|\alpha)$ with respect to σ^p and $\tilde{\theta}$.

Theorem 4. *Let* $(\sigma^p,\tilde{\theta}) \in \Xi^p \times \Theta$. *Then, for a given nominal time-delay* α *and each* $\epsilon > 0$ *and* $\gamma > 0$,

$$
\frac{\partial H^{\epsilon,\gamma}(\sigma^p,\tilde{\theta}|\alpha)}{\partial \tilde{\theta}_i} = \tilde{\lambda}^{\top}(\tilde{\theta}_i)\hat{f}^{i,n_{p_i}}(\tilde{\theta}_i|\sigma^p,\tilde{\theta},\alpha) - \tilde{\lambda}^{\top}(\tilde{\theta}_i)\hat{f}^{i+1,1}(\tilde{\theta}_i|\sigma^p,\tilde{\theta},\alpha),
$$

$$
i = 1,\ldots,N-1,
$$

$$
\frac{\partial H^{\epsilon,\gamma}(\sigma^p,\tilde{\theta}|\alpha)}{\partial \sigma^{p_i,l}} = \int_{\theta_{l-1}^{p_i}}^{\theta_l^{p_i}}\tilde{\lambda}^{\top}(s)\frac{\partial \hat{f}^{i,l}(s|\sigma^p,\tilde{\theta},\alpha)}{\partial \sigma^{p_i,l}}ds, \quad l = 1,\ldots,n_{p_i},\ i = 1,\ldots,N,
$$

and

$$\frac{\partial H^{\epsilon,\gamma}(\sigma^p,\tilde{\theta}|\alpha)}{\partial t_f} = \sum_{i=1}^{N}\sum_{l=1}^{n_{p_i}}\int_{\theta_{l-1}^{p_i}}^{\theta_l^{p_i}} t_f^{-1}\tilde{\lambda}^{\top}(s)\hat{f}^{i,l}(s|\sigma^p,\tilde{\theta},\alpha)ds$$

$$+ \sum_{i=1}^{N}\sum_{l=1}^{n_{p_i}}\int_{\theta_{l-1}^{p_i}}^{\theta_l^{p_i}} t_f^{-2}\alpha\tilde{\lambda}^{\top}(s)\frac{\partial \hat{f}^{i,k}(s|\sigma^p,\tilde{\theta},\alpha)}{\partial \hat{x}}\check{\zeta}^p(s - t_f^{-1}\alpha|\sigma^p,\tilde{\theta},\alpha)ds$$

$$+ \sum_{i=1}^{N}\sum_{l=1}^{n_{p_i}}\int_{\theta_{l-1}^{p_i}-t_f^{-1}\alpha}^{\theta_l^{p_i}-t_f^{-1}\alpha} \left\{ s\tilde{\lambda}^{\top}(s + t_f^{-1}\alpha)\frac{\partial \hat{f}^{i,l}(s + t_f^{-1}\alpha|\sigma^p,\tilde{\theta},\alpha)}{\partial \tilde{x}} \right.$$

$$\left. \times \dot{\phi}(t_f s)\chi_{(-\infty,0]}(s) \right\}ds,$$

where $\tilde{\lambda}(\cdot) := \tilde{\lambda}(\cdot|\sigma^p,\tilde{\theta},\alpha)$ *is the solution of the following costate system:*

$$\dot{\tilde{\lambda}}(s) = -\sum_{i=1}^{12}\frac{\partial \pi_{\epsilon}(h_i(\tilde{x}^p(s|\sigma^p,\tilde{\theta},\alpha)))}{\partial \tilde{x}} - \left(\sum_{i=1}^{N}\sum_{l=1}^{n_{p_i}}\frac{\partial \hat{f}^{i,l}(s|\sigma^p,\tilde{\theta},\alpha)}{\partial \tilde{x}}\right)^{\top}$$

$$\times \tilde{\lambda}(s)\chi_{\mathcal{S}_l^{p_i}(\theta,t_f,0)}(s) - \left(\sum_{i=1}^{N}\sum_{l=1}^{n_{p_i}}\frac{\partial \hat{f}^{i,l}(s + t_f^{-1}\alpha|\sigma^p,\tilde{\theta},\alpha)}{\partial \tilde{x}}\right)^{\top}$$

$$\times \tilde{\lambda}(s + t_f^{-1}\alpha)\chi_{\mathcal{S}_l^{p_i}(\theta,t_f,\alpha)}(s), \qquad\qquad s \in [0,1],$$

with the terminal condition

$$\tilde{\lambda}(s) = (0,0,0,0,0,0)^{\top}, \quad s \geq 1.$$

Based on Theorems 2-4, the gradients of constraints (39) and (42) with respect to σ^p and $\tilde{\theta}$ can be obtained. Moreover, the gradients of the objective (38) and the constraint (39) with respect to ν can be easily computed. Note that these gradients can be integrated with a standard nonlinear optimization method— e.g., SQP—to solve each of the nonlinear programming problems obtained through the application of the control parameterization in conjunction with the constraint transcription technique.

6. Numerical results

In the numerical simulation, we consider a 1,3-PD fed-batch production process by *K. pneumoniae* reported in [52]. The maximal duration of this process is partitioned into the first batch phase (Ph. I) and phases II-X (Phs. II-X) according to the number of switchings. The same time durations and feeding rates of glycerol for feeding processes are adopted in each one of Phs. II-X. Moreover, within each one of Phs. II-X, all batch processes have 100s minus the duration of the feeding process. The characteristic of each phase are given in Table 2. As a result, the terminal time, the end moment of Ph. I, the feeding rate of glycerol and the end moments of the first feeding process in Phs. II-X should be optimized.

Table 2: Phase characteristics in fed-batch process [52]. Each phase consists of a series of identical 100s feeding batch combinations. Within each phase, all feeding modes have the same duration and all batch modes have the same duration.

Phase	Start time (h)	End time (h)	Number of modes		Mode duration (s)	
			Feeding	Batch	Feeding	Batch
I	0	5.3300	0	1	0	19188
II	5.3300	6.1078	28	28	5	95
III	6.1078	7.1356	37	37	7	93
IV	7.1356	8.8300	61	61	8	92
V	8.8300	12.1356	119	119	7	93
VI	12.1356	15.8300	133	133	6	94
VII	15.8300	18.0800	81	81	4	96
VIII	18.0800	19.8300	63	63	3	97
IX	19.8300	23.8300	144	144	2	98
X	23.8300	24.1633	12	12	1	99

To numerically solve the required SOC problems, we wrote a Fortran subroutine, in which the control parameterization, the constraint transcription and the gradients in the previous Section 5.2 are combined with the optimization software NLPQLP—a Fortran implementation of SQP [53]. This subroutine uses the differential equation software DLSODA [54] to solve the state and costate systems. Lagrange interpolation [55] is used whenever the DLSODA requires the value of the state or costate at an intermediate time between two adjacent knot points. The nominal time-delay, the history function, and the number of partitions are $\alpha = 0.4652$h, $\phi(t) = (0.1115\text{gL}^{-1}, 495\text{mmolL}^{-1}, 0, 0, 0, 5\text{L})^\top$, and $n_{p_i} = 4$, respectively [19]. Note that there are 10 phases and 1355 switchings between batch and feeding modes in the maximal duration of the fed-batch process [52]. For this situation, we denote the end moment of the first batch mode by τ_1, and the end moment of the first feeding mode in each phase of Phs. II-X by τ_i, $i = 2, \ldots, 10$. The initial values of switching instants and terminal time [56] are listed in Table 3. It should be noted that these initial values of switching instants are obtained from Table 2. In addition, since the given initial value of the terminal time, 14.16h, is greater than the start time of Ph. VI and is less than the end time of Ph. VI (see Table 2), only six switching instants, $\tau_1, \ldots, \tau_6$, are to be optimized in the first iteration of SOC problems. Therefore, we just listed these six values and the other values of switching instants are missing entries (denoted by hyphens). The lower and upper bound vectors in (7) are given as $\rho = (5.3, 0.0006, 0.6767, 0.9261, 1.5928, 3.2039, 3.5919, 2.1483, 1.6483, 3.8983, 11)^\top$ and $\delta = (5.4, 0.0022, 0.88, 1.129, 1.7961, 3.4072, 3.7953, 2.3517, 1.8517, 4.1006, 24.1633)^\top$, respectively [56]. From vectors ρ and δ, we obtained the lower and upper bounds of switching instants and terminal time, which are also listed in Table 3. Similarly, the missing entries in the column of the lower bound in Table 3 denote that switchings occur after the terminal time. Besides, the

Table 3: Initial values, optimal values, and lower and upper bounds for the switching instants and terminal time [56]. Note that the missing entries in the table denote that the switchings occur after the terminal time.

	Initial value	Lower bound	Upper bound	Weight $\omega = (0,1)^\top$		Weight $\omega = (1,0)^\top$	
				$\beta = 0$	$\beta = 1.0$	$\beta = 0$	$\beta = 1.0$
τ_1	5.3300	5.3000	5.4000	5.3000	5.3000	5.3000	5.3000
τ_2	5.3314	5.3006	5.4022	5.3015	5.3010	5.3005	5.3006
τ_3	6.1097	6.0789	6.1806	6.0805	6.0796	6.0788	6.07956
τ_4	7.1378	7.1067	7.2083	7.1075	7.1079	7.1082	7.1083
τ_5	8.8319	8.8011	8.9028	8.8011	8.8015	8.8027	8.8013
τ_6	12.1372	-	12.2083	12.1067	12.1066	12.1083	12.1067
τ_7	-	-	15.9027	15.9019	15.8003	-	-
τ_8	-	-	18.1519	18.1519	18.0503	-	-
τ_9	-	-	19.9019	-	-	-	-
τ_{10}	-	-	23.9008	-	-	-	-
t_f	14.1600	11.0000	24.1633	18.1611	19.3228	14.0867	15.4944

Table 4: Initial values and bounds of glycerol feeding rates in phases II-X [13].

Phases	II-III	IV	V-VI	VII	VIII	IX-X
Initial value (mLs^{-1})	0.2103	0.1992	0.2103	0.2214	0.2437	0.2548
Lower bound (mLs^{-1})	0.1682	0.1594	0.1682	0.1771	0.1949	0.2038
Upper bound (mLs^{-1})	0.2524	0.2390	0.2524	0.2657	0.2924	0.3058

initial values and bounds of glycerol feeding rates in Phs. II-X [13] are listed in Table 4. Furthermore, in the constrain transcription, we take $\epsilon = 1.0 \times 10^{-2}$ and $\gamma = 2.5 \times 10^{-3}$ as the initial values for the adjustable parameters, adjusting them at each step according to $\epsilon - \gamma$ process [37]. This process involves reducing γ by a factor of 2 if the optimal solution satisfies (36), or reducing both ϵ and γ by a factor of 10 if the solution does not satisfy (36). We terminated the ϵ–γ process when $\epsilon \leq 1.0 \times 10^{-8}$ and $\gamma \leq 1.0 \times 10^{-7}$.

To solve (RMOP$^\beta$) for a given weight β by using the modified NBI method, we wrote a Fortran program. This program combines the modified NBI with the single-objective optimal control solver described above using evenly distributed weights, in which the uniform spacing between two consecutive ω_1 is $1/10$ and $\omega_2 = 1 - \omega_1$ (resulting in 11 SOC problems). By respectively running this program for $\beta = 0$, $\beta = 0.5$, and $\beta = 1.0$, we obtain points in the Pareto sets as depicted in Figure 1. Moreover, we solve (RMOP$^\beta$) for $\beta = 1.0$ by using the original NBI method and the CWS method, in which two components of J^β are assigned weights and then added together to form a single objective, with
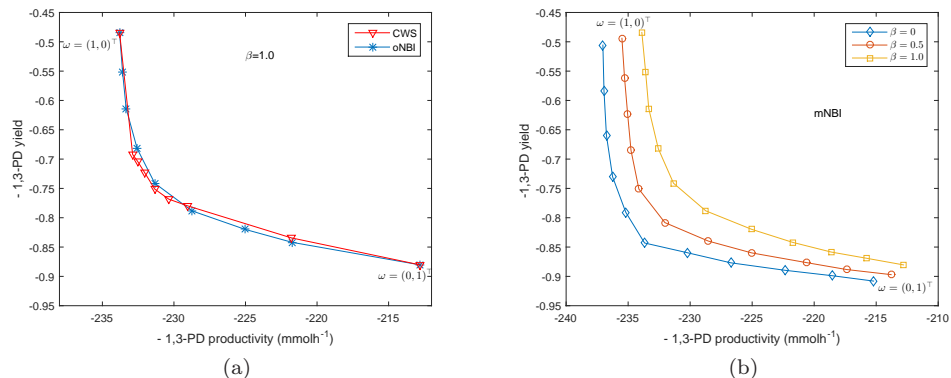
Figure 1: (a). Pareto sets generated by the original NBI and the CWS for $\beta = 1.0$; and (b) Pareto sets generated by the modified NBI methods for $\beta = 0, 0.5, 1.0$.

the same weight ω. The obtained points in the Pareto sets are also plotted in Figure 1. As expected, minimizing the negative value of 1,3-PD productivity and minimizing the negative value of 1,3-PD yield are conflicting because an increase of one results in lowering the other. Although all the MOC approaches capture this feature, large differences in the accuracy of the resulting Pareto sets are visible. For example, for $\beta = 1.0$, although a uniform grid is used to vary the weights in all three MOC approaches, the results for the modified NBI clearly exhibit more evenly distributed points along the Pareto set than the ones for the original NBI and the CWS. Both the original NBI and the CWS generate non-global Pareto points, which have been eliminated by the filter designed in Subsection 5.1, and hence, the modified NBI provides a more accurate representation of the Pareto set.

For the Pareto points obtained by using modified NBI method (see Figure 1(b)), the corresponding optimal switching instants and terminal times are also listed in Table 3. Note that the missing entries in the columns of the optimal switching instants and terminal times in Table 3 also denote that switchings occur after the terminal times. Figure 2 illustrates the corresponding optimal feeding rates of glycerol and the corresponding optimal feeding durations. It should be noted that Ph. I is a batch process and no glycerol is fed in this phase. Ph. VII and Ph. VIII only include the feeding rates of glycerol and the feeding durations for points with weight $\omega = (0,1)^\top$ and have much shorter feeding durations than Phs. II-VI. As can be seen from Table 3 and Figure 2, the optimal terminal times for points with weight $\omega = (0,1)^\top$ are greater than the optimal terminal times for points with weight $\omega = (1,0)^\top$. This observation is explained by the fact: for points with weight $\omega = (0,1)^\top$, only the negative value of 1,3-PD yield is minimized, which allows longer process times. In contrast, for points with weight $\omega = (1,0)^\top$, the negative value of 1,3-productivity is minimized, in which shorter process times are needed. Under these optimal switching control strategies, the changes of 1,3-PD and the volume of culture
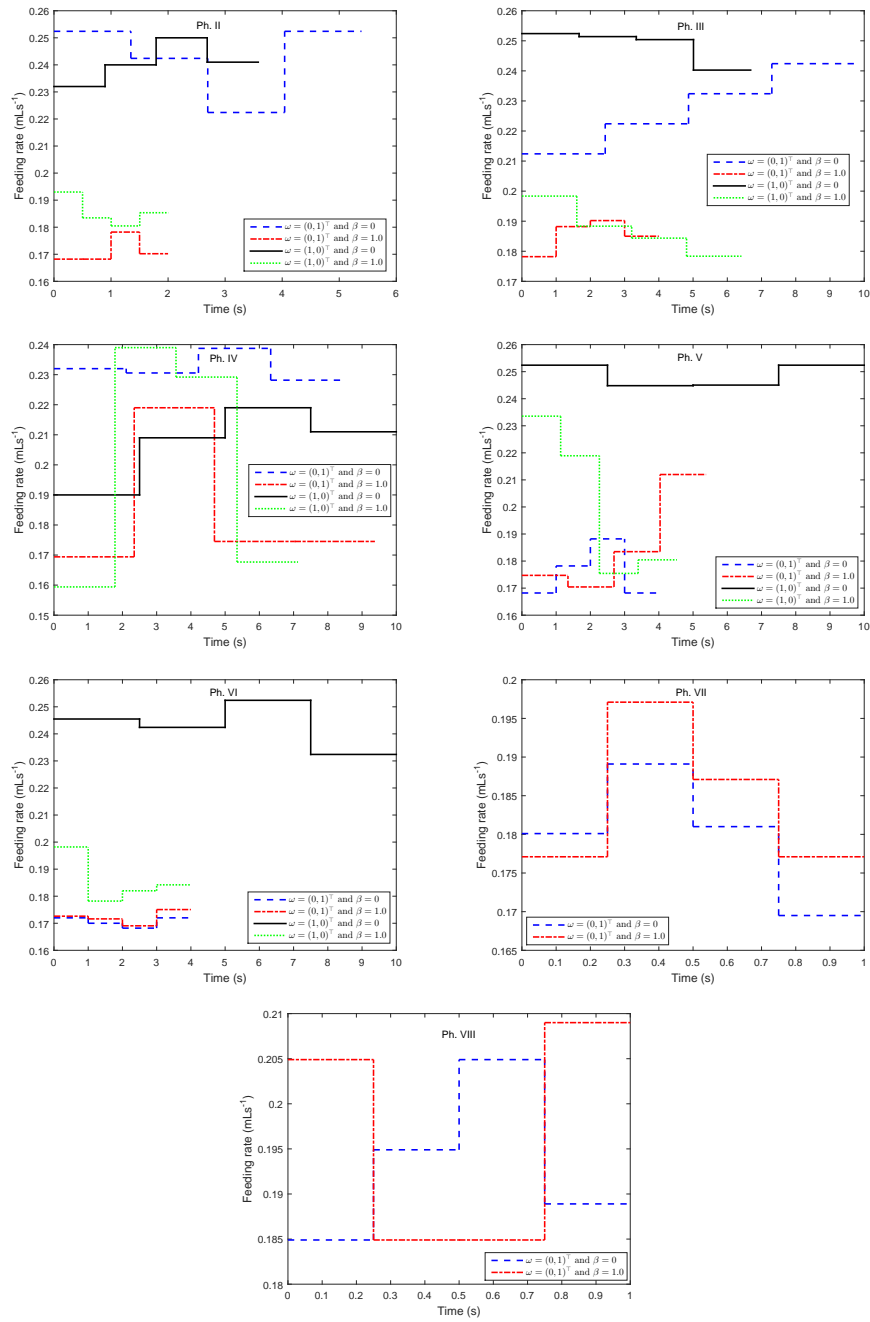
24

Figure 2: The feeding rates of glycerol and the feeding durations in the fermentation process.
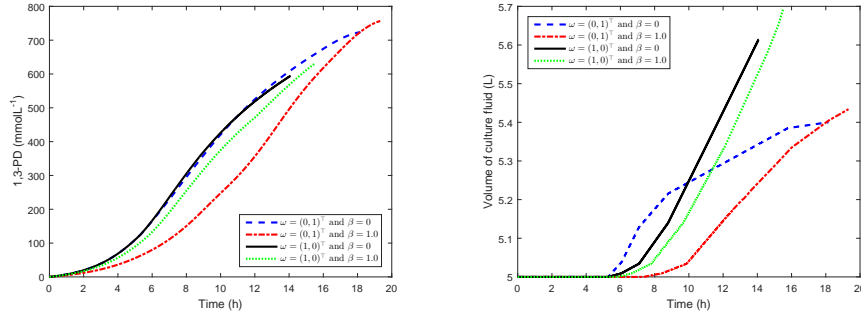
Figure 3: The changes of 1,3-PD concentration and volume of culture fluid with respect to fermentation time.

fluid with respect to the fermentation time are plotted in Figure 3. As can be seen from Figure 3, less 1,3-PD concentrations are attained for points with weight $\omega = (1,0)^\top$ than the ones for points with weight $\omega = (0,1)^\top$. In particular, for $\beta = 0, 1.0$, the attained 1,3-PD concentrations are $(594.8071, 636.2763)^\top$ for points with weight $\omega = (1,0)^\top$ compared with $(723.7839, 756.6823)^\top$ for points with weight $\omega = (0,1)^\top$. However, more substrates are fed within shorter process times for points with weight $\omega = (1,0)^\top$ than the ones for points with weight $\omega = (0,1)^\top$. Specifically, for $\beta = 0, 1.0$, the final volumes of culture fluid are $(5.6137, 5.6958)^\top$ for points with weight $\omega = (1,0)^\top$ compared with $(5.4, 5.434)^\top$ for points with weight $\omega = (0,1)^\top$. Again, these attained 1,3-PD concentrations and final volumes of culture fluid confirm the confliction of 1,3-PD productivity and 1,3-PD yield.

From Figure 1(b), we can also see that when the value of β increases, i.e, the robustness of 1,3-PD productivity and 1,3-PD yield with respect to the uncertain time-delay is emphasized on, 1,3-PD productivity and 1,3-PD yield are correspondingly degraded. For example, for the points with weight $\omega = (1,0)^\top$ in the Pareto sets, 1,3-PD productivity and 1,3-PD yield decrease from $(237.0349, 0.5056)^\top$ for $\beta = 0$ to $(233.8959, 0.484)^\top$ for $\beta = 1.0$. Similarly, for points with weight $\omega = (0,1)^\top$ in the Pareto sets, 1,3-PD productivity and 1,3-PD yield decrease from $(215.2076, 0.908)^\top$ for $\beta = 0$ to $(212.7946, 0.8804)^\top$ for $\beta = 1.0$. This implies that the optimal switching control of (RMOP^β) for $\beta = 0$ is more favorable than the one for $\beta = 1.0$ when the time-delay α can be exactly obtained. On the other hand, when there is uncertainty in the time-delay, we investigated the disturbance of the nominal time-delay α for points with weights $\omega = (1,0)^\top$ and $\omega = (0,1)^\top$ in the Pareto sets. We randomly generated 50 perturbations of the nominal time-delay α. For each $i = 1, \ldots, 50$, the perturbed time-delay belong to the set of all $\bar{\alpha}$ such that

$$0.01(i-1)\alpha \leq |\bar{\alpha} - \alpha| \leq 0.01i\alpha.$$

The coefficient of α in the upper bound of this range is called the disturbance percentage of the perturbed time-delay. With the optimal switching controls,
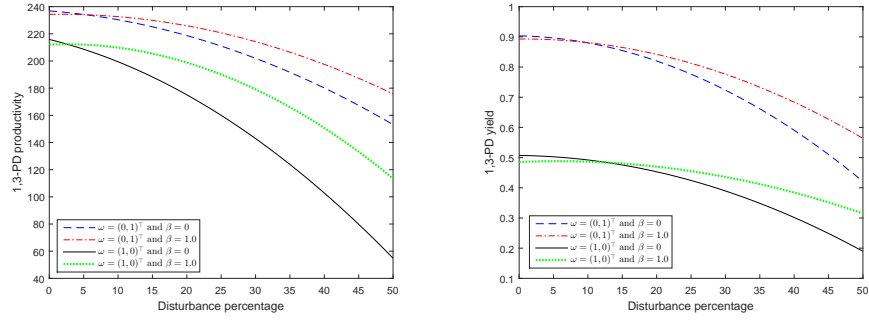
26

Figure 4: 1,3-PD productivity and 1,3-PD yield variation curves under time-delay disturbance.

the variations of 1,3-PD productivity and 1,3-PD yield under the 50 randomly-selected perturbed time-delays are shown in Figure 4. Note that the variation curves for $\beta = 1.0$ predominantly lie above the variation curves for $\beta = 0$. This indicates that the optimal switching control obtained by solving (RMOP^{β}) for $\beta = 1.0$ is more robust than the one for $\beta = 0$, especially for large values of the disturbance percentage. Thus, the optimal switch control of (RMOP^{β}) for $\beta = 1.0$ is more preferable than the one for $\beta = 0$ when there is uncertainty in the time-delay.

7. Conclusion

This paper has considered a robust multi-objective optimal control of nonlinear switched time-delay system arising in 1,3-PD fed-batch production process. The aim of the problem is to design an optimal switching control strategy such that 1,3-PD productivity and 1,3-PD yield are maximized and their sensitivities with respect to uncertain time-delay are minimized, simultaneously. The robust multi-objective optimal control problem is transformed into an equivalent one in standard form. A numerical solution method combining with a modified normal boundary intersection method and a single-objective solver is developed to solve the equivalent problem. Numerical results verified the effectiveness of the proposed solution method. In closing, we note that there may be other uncertainties in the fed-batch process, for example, uncertainties in kinetic parameter and initial state values. Thus, designing optimal switching control scheme in the presence of these uncertainties is an interesting area to pursue for future research.

Acknowledgments

Appendix

The definitions of $\mathcal{L}_1$, $\mathcal{L}_2$, $\mathcal{L}_3$ and $\mathcal{L}_4$ in Theorems 2 and 3 are given below.

$$
\begin{aligned}
\mathcal{L}_1(s) := {}& t_f^{-1} \frac{\partial \hat{f}^{i,l}(s|\sigma^p, \tilde{\theta}, \alpha)}{\partial \tilde{x}} \tilde{\varphi}^p(s) + t_f^{-1} \frac{\partial \hat{f}^{i,l}(s|\sigma^p, \tilde{\theta}, \alpha)}{\partial \hat{x}} \tilde{\varphi}^p(s - t_f^{-1}\alpha) \\
& + t_f^{-2}\alpha \frac{\partial^2 \hat{f}^{i,l}(s|\sigma^p, \tilde{\theta}, \alpha)}{\partial \hat{x} \partial \tilde{x}} \tilde{\zeta}^p(s - t_f^{-1}\alpha|\sigma^p, \tilde{\theta}, \alpha)\tilde{\varphi}^p(s) \\
& + t_f^{-2}\alpha \frac{\partial \hat{f}^{i,l}(s|\sigma^p, \tilde{\theta}, \alpha)}{\partial \hat{x}} \psi(s - t_f^{-1}\alpha|\sigma^p, \tilde{\theta}, \alpha) \\
& + t_f^{-2}\alpha \frac{\partial^2 \hat{f}^{i,l}(s|\sigma^p, \tilde{\theta}, \alpha)}{\partial \hat{x} \partial \tilde{x}} \tilde{\zeta}^p(s - t_f^{-1}\alpha|\sigma^p, \tilde{\theta}, \alpha)\tilde{\varphi}^p(s - t_f^{-1}\alpha) \\
& - t_f^{-3}\alpha \frac{\partial^2 \hat{f}^{i,l}(s|\sigma^p, \tilde{\theta}, \alpha)}{\partial \hat{x}^2} \tilde{\zeta}^p(s - t_f^{-1}\alpha|\sigma^p, \tilde{\theta}, \alpha)^2 \\
& - t_f^{-3}\alpha \frac{\partial \hat{f}^{i,l}(s|\sigma^p, \tilde{\theta}, \alpha)}{\partial \hat{x}} \varpi(s - t_f^{-1}\alpha|\sigma^p, \tilde{\theta}, \alpha),
\end{aligned} \tag{A.1}
$$

$$
\mathcal{L}_2(s) := t_f^{-1}\hat{f}^{i,l}(s|\sigma^p, \tilde{\theta}, \alpha) + t_f^{-2}\alpha \frac{\partial \hat{f}^{i,l}(s|\sigma^p, \tilde{\theta}, \alpha)}{\partial \hat{x}} \tilde{\zeta}^p(s - t_f^{-1}\alpha|\sigma^p, \tilde{\theta}, \alpha), \tag{A.2}
$$

$$
\begin{aligned}
\mathcal{L}_3(s) := {}& \frac{\partial^2 \hat{f}^{i,l}(s + t_f^{-1}\alpha|\sigma^p, \tilde{\theta}, \alpha)}{\partial \tilde{x} \partial \bar{x}(s)} \tilde{\varphi}^p(s + t_f^{-1}\alpha) \\
& + \frac{\partial^2 \hat{f}^{i,l}(s + t_f^{-1}\alpha|\sigma^p, \tilde{\theta}, \alpha)}{\partial \tilde{x}^2} \tilde{\varphi}^p(s) \\
& - t_f^{-1} \frac{\partial}{\partial \tilde{x}} \left\{ \frac{\partial \hat{f}^{i,l}(s + t_f^{-1}\alpha|\sigma^p, \tilde{\theta}, \alpha)}{\partial \tilde{x}} \tilde{\zeta}^p(s|\sigma^p, \tilde{\theta}, \alpha) \right\},
\end{aligned} \tag{A.3}
$$

and

$$
\mathcal{L}_4(s) := \frac{\partial \hat{f}^{i,l}(s + t_f^{-1}\alpha|\sigma^p, \tilde{\theta}, \alpha)}{\partial \tilde{x}}, \tag{A.4}
$$

where $\tilde{\varphi}^p(\cdot) := \tilde{\varphi}^p(\cdot|\sigma^p, \tilde{\theta}, \alpha)$; and $\tilde{\zeta}^p(\cdot|\sigma^p, \tilde{\theta}, \alpha)$, $\hat{f}^{i,l}(\cdot|\sigma^p, \tilde{\theta}, \alpha)$, $\psi(\cdot|\sigma^p, \tilde{\theta}, \alpha)$ and $\varpi(\cdot|\sigma^p, \tilde{\theta}, \alpha)$ are as defined in (35), (43), (44) and (45), respectively.

The explicit formulas for $\partial \tilde{J}_1^{\beta,p}/\partial \tilde{\varphi}_3$, $\partial \tilde{J}_1^{\beta,p}/\partial \tilde{\varphi}_6$, $\partial \tilde{J}_1^{\beta,p}/\partial \tilde{x}_3$ and $\partial \tilde{J}_1^{\beta,p}/\partial \tilde{x}_6$

in Theorem 2 are given below.

$$\frac{\partial \tilde{J}_1^{\beta,p}(\sigma^p, \tilde{\theta}|\alpha)}{\partial \tilde{\varphi}_3} = 2\beta t_f^{-2} \tilde{x}_6^p(1) \big[\tilde{\varphi}_3^p(1)\tilde{x}_6^p(1) + \tilde{\varphi}_6^p(1)\tilde{x}_3^p(1) \big], \tag{A.5}$$

$$\frac{\partial \tilde{J}_1^{\beta,p}(\sigma^p, \tilde{\theta}|\alpha)}{\partial \tilde{\varphi}_6} = 2\beta t_f^{-2} \tilde{x}_3^p(1) \big[\tilde{\varphi}_3^p(1)\tilde{x}_6^p(1) + \tilde{\varphi}_6^p(1)\tilde{x}_3^p(1) \big], \tag{A.6}$$

$$\frac{\partial \tilde{J}_1^{\beta,p}(\sigma^p, \tilde{\theta}|\alpha)}{\partial \tilde{x}_3} = 2\beta t_f^{-2} \tilde{\varphi}_6^p(1) \big[\tilde{\varphi}_3^p(1)\tilde{x}_6^p(1) + \tilde{\varphi}_6^p(1)\tilde{x}_3^p(1) \big] - t_f^{-1}\tilde{x}_6^p(1), \tag{A.7}$$

$$\frac{\partial \tilde{J}_1^{\beta,p}(\sigma^p, \tilde{\theta}|\alpha)}{\partial \tilde{x}_6} = 2\beta t_f^{-2} \tilde{\varphi}_3^p(1) \big[\tilde{\varphi}_3^p(1)\tilde{x}_6^p(1) + \tilde{\varphi}_6^p(1)\tilde{x}_3^p(1) \big] - t_f^{-1}\tilde{x}_3^p(1), \tag{A.8}$$

where $\tilde{x}^p(1) := \tilde{x}^p(1|\sigma^p, \tilde{\theta}, \alpha)$; and $\tilde{\varphi}^p(1) := \tilde{\varphi}^p(1|\sigma^p, \tilde{\theta}, \alpha)$.

The explicit formulas for $\partial \tilde{J}_2^{\beta,p}/\partial \tilde{\varphi}_3$, $\partial \tilde{J}_2^{\beta,p}/\partial \tilde{\varphi}_6$, $\partial \tilde{J}_2^{\beta,p}/\partial \tilde{x}_3$ and $\partial \tilde{J}_2^{\beta,p}/\partial \tilde{x}_6$ in Theorem 3 are given below.

$$\begin{aligned}\frac{\partial \tilde{J}_2^{\beta,p}(\sigma^p, \tilde{\theta}|\alpha)}{\partial \tilde{\varphi}_3} &= 2\beta c_{s0}^{-2} \tilde{x}_6^p(1)\big[\tilde{\varphi}_3^p(1)\tilde{x}_6^p(1)^2 - \tilde{\phi}_6(0)\tilde{\varphi}_3^p(1)\tilde{x}_6^p(1) \\ &\quad - \tilde{\phi}_6(0)\tilde{\varphi}_6^p(1)\tilde{x}_3^p(1) \big]\big[\tilde{x}_6^p(1) - \tilde{\phi}_6(0) \big]^{-3}, \end{aligned} \tag{A.9}$$

$$\begin{aligned}\frac{\partial \tilde{J}_2^{\beta,p}(\sigma^p, \tilde{\theta}|\alpha)}{\partial \tilde{\varphi}_6} &= -2\beta c_{s0}^{-2}\tilde{\phi}_6(0)\tilde{x}_3^p(1)\big[\tilde{\varphi}_3^p(1)\tilde{x}_6^p(1)^2 - \tilde{\phi}_6(0)\tilde{\varphi}_3^p(1)\tilde{x}_6^p(1) \\ &\quad - \tilde{\phi}_6(0)\tilde{\varphi}_6^p(1)\tilde{x}_3^p(1) \big]\big[\tilde{x}_6^p(1) - \tilde{\phi}_6(0) \big]^{-4}, \end{aligned} \tag{A.10}$$

$$\begin{aligned}\frac{\partial \tilde{J}_2^{\beta,p}(\sigma^p, \tilde{\theta}|\alpha)}{\partial \tilde{x}_3} &= -c_{s0}^{-1}\tilde{x}_6^p(1)\big[\tilde{x}_6^p(1) - \tilde{\phi}_6(0) \big]^{-1} - 2\beta c_{s0}^{-2}\tilde{\phi}_6(0)\tilde{\varphi}_6^p(1) \\ &\quad \times \big[\tilde{\varphi}_3^p(1)\tilde{x}_6^p(1)^2 - \tilde{\phi}_6(0)\tilde{\varphi}_3^p(1)\tilde{x}_6^p(1) - \tilde{\phi}_6(0)\tilde{\varphi}_6^p(1)\tilde{x}_3^p(1) \big] \\ &\quad \times \big[\tilde{x}_6^p(1) - \tilde{\phi}_6(0) \big]^{-4}, \end{aligned} \tag{A.11}$$

and

$$\begin{aligned}\frac{\partial \tilde{J}_2^{\beta,p}(\sigma^p, \tilde{\theta}|\alpha)}{\partial \tilde{x}_6} &= c_{s0}^{-1}\tilde{\phi}_6(0)\tilde{x}_3^p(1)\big[\tilde{x}_6^p(1) - \tilde{\phi}_6(0) \big]^{-2} - 2\beta c_{s0}^{-2}\tilde{\phi}_6(0)\big[\tilde{\phi}_6(0)\tilde{\varphi}_6^p(1)\tilde{x}_3^p(1) \\ &\quad - \tilde{\varphi}_3^p(1)\tilde{x}_6^p(1)^2 + \tilde{\phi}_6(0)\tilde{\varphi}_3^p(1)\tilde{x}_6^p(1) \big]\big[2\tilde{\varphi}_6^p(1)\tilde{x}_3^p(1) - \tilde{\varphi}_3^p(1)\tilde{x}_6^p(1) \\ &\quad + \tilde{\phi}_6(0)\tilde{\varphi}_3^p(1) \big]\big[\tilde{x}_6^p(1) - \tilde{\phi}_6(0) \big]^{-5}, \end{aligned} \tag{A.12}$$

where $\tilde{x}^p(1) := \tilde{x}^p(1|\sigma^p, \tilde{\theta}, \alpha)$; and $\tilde{\varphi}^p(1) := \tilde{\varphi}^p(1|\sigma^p, \tilde{\theta}, \alpha)$.

References

[1] V.S. Bisaria, A. Kondo, Bioprocessing of Renewable Resources to Commodity Bioproducts, John Wiley & Sons Inc., New Jersey, 2014.

[2] H. Liu, Y. Xu, Z. Zheng, D. Liu, 1,3-Propanediol and its copolymers: Research, development and industrialization, Biotechnol. J., **5** (2010), 1137–1148.

[3] J. Chaminand, L. Djakovitch, P. Gallezot, P. Marion, C. Pinel, C. Rosier, Glycerol hydrogenolysis on heterogeneous catalysts, Green Chem., **6** (2004), 359–361.

[4] G. Aggelis, Microbial Conversions of Raw Glycerol, Nova Science Publishers, New York, 2009.

[5] K. Menzel, A. Zeng, W. Deckwer, High concentration and productivity of 1,3-propanediol from continuous fermentation of glycerol by *Klebsiella pneumoniae*, Enzyme Microb. Technol., **20** (1997), 82–86.

[6] A. Zeng, H. Biebl, Bulk chemicals from biotechnology: The case of 1,3-propanediol production and the new trends, Adv. Biochem. Eng. Biotechnol., **74** (2002), 239–259.

[7] A. Mészáros, V. Bales, A contribution to optimal control of fed-batch biochemical processes, Bioproc. Biosyst. Eng., **7** (1992), 363–367.

[8] J. Lee, S. Lee, S. Park, A.P.J. Middelberg, Control of fed-batch fermentations, Biotechnol. Adv., **17** (1999), 29–48.

[9] A. Ashoori, B. Moshiri, A. Khaki-Sedigh, M.R. Bakhtiari, Optimal control of a nonlinear fed-batch fermentation process using model predictive approach, J. Process Contr., **19** (2009), 1162–1173.

[10] H. Wang, E. Feng, Z. Xiu, Optimality condition of the nonlinear impulsive system in fed-batch fermentation, Nonlinear Anal. Theor., **68** (2008), 12–23.

[11] J. Gao, B. Shen, E. Feng, Z. Xiu, Modelling and optimal control for an impulsive dynamical system in microbial fed-batch culture, Comput. Appl. Math., **32** (2013), 275–290.

[12] C. Liu, Z. Gong, E. Feng, H. Yin, Optimal switching control for microbial fed-batch culture, Nonlinear Anal. Hybrid Syst., **2** (2008), 1168–1174.

[13] C. Liu, Z. Gong, E. Feng, H. Yin, Modelling and optimal control for nonlinear multistage dynamical system of microbial fed-batch culture, J. Ind. Manag. Optim., **5** (2009), 835–850.

[14] L. Wang, Q. Lin, R. Loxton, K.L. Teo, G. Cheng, Optimal 1,3-propanediol production: Exploring the trade-off between process yield and feeding rate variation, J. Process Contr., **32** (2015), 1–9.

[15] K. Menzel, A.P. Zeng, H. Biebl, W.D. Deckwer, Kinetic, dynamic, and pathway studies of glycerol metabolism by *Klebsiella pneumoniae* in anaerobic continuous culture: I. The phenomena and characterization of oscillation and hysteresis, Biotechnol. Bioeng., **52** (1996), 549–560.

[16] Z. Xiu, B. Song, L. Sun, A. Zeng, Theoretical analysis of effects of metabolic overflow and time delay on the performance and dynamic behavior of a two-stage fermentation process, Biochem. Eng. J., **11** (2002), 101–109.

[17] N. MacDonald, Time-lags in Biological Models, Lect. Notes Biomath., **27**, Springer, Berlin, 1979.

[18] U. an der Heiden, Delays in physiological systems, J. Math. Biol., **8** (1979), 345–364.

[19] C. Liu, Sensitivity analysis and parameter identification for a nonlinear time-delay system in microbial fed-batch process, Appl. Math. Model., **38** (2014), 1449–1463.

[20] C. Liu, Z. Gong, Optimal Control of Switched Systems Arising in Fermentation Processes, Springer-Verlag, Berlin, 2014.

[21] A. Giua, C. Seatze, C. Van Der Mee, Optimal control of autonomous linear systems switched with a preassigned finite sequence, In: Proceedings of 2001 IEEE ISIC, México City, México, 2001, pp. 144–149.

[22] X. Xu, P.J. Antsaklis, Optimal control of switched systems based on parameterization of the switching instants, IEEE Trans. Autom. Control, **49** (2004), 2–15.

[23] S. Bengea, A.D. Raymond, Optimal control of switching systems, Automatica, **41** (2005), 11–27.

[24] Q. Lin, R. Loxton, K.L. Teo, Optimal control of nonlinear switched systems: Computational methods and applications, J. Oper. Res. Soc. China, **1** (2013), 275–311.

[25] F. Delmotte, E.I. Verriest, M. Egersedt, Optimal impulsive control of delay systems, ESAIM Control Optim. Calc. Var., **14** (2008), 767–779.

[26] C. Liu, R. Loxton, K.L. Teo, Switching time and parameter optimization in nonlinear switched systems with multiple time-delays, J. Optim. Theory Appl., **163** (2014), 957–988.

[27] F. Jadot, G. Bastin, J.F. Van Impe, Optimal adaptive control of a bioprocess with yield-productivity conflict, J. Biotechnol., **65** (1998), 61–68.

[28] V. Bhaskar, S. Gupta, A. Ray, Applications of multi-objective optimization in chemical engineering, Rev. Chem. Eng., **16** (2000), 1–54.

[29] D. Sarkar, J.M. Modak, Pareto-optimal solutions for multi-objective optimization of fed-batch bioreactors using nondominated sorting genetic algorithm, Chem. Eng. Sci., **60** (2005), 481–492.

[30] K. Muske, J. Peyton, Multi-objective model-based control for an automotive catalyst, J. Process Contr., **16** (2006), 27–35.

[31] Y. Lim, P. Floquet, X. Joulia, Efficient implementation of the normal boundary intersection (NBI) method on multiobjective optimization problems, Ind. Eng. Chem. Res., **40** (2001), 648–655.

[32] N. Srinivas, K. Deb, Multiobjective function optimization using nondominated sorting genetic algorithms, Evol. Comput. J., **2** (1995), 221–248.

[33] C.A.C. Coello, G.T. Pulido, M.S. Lechuga, Handling multiple objectives with particle swarm optimization, IEEE Trans. Evol. Comput., **8** (2004), 256–279.

[34] F. Logist, P. Van Erdeghem, J. Van Impe, Efficient deterministic multiple objective optimal control of (bio)chemical processes, Chem. Eng. Sci., **64** (2009), 2527–2538.

[35] F. Logist, B. Houska, M. Diehl, J. Van Impe, Robust multi-objective optimal control of uncertain (bio)chemical processes, Chem. Eng. Sci., **66** (2011), 4670–4682.

[36] C. Liu, R. Loxton, K.L. Teo, A computational method for solving time-delay optimal control problems with free terminal time, Syst. Control Lett., **72** (2014), 53–60.

[37] 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.

[38] K.L. Teo, C.J. Goh, K.H. Wong, A Unified Computational Approach to Optimal Control Problems, Longman Scientific and Technical, Essex, 1991.

[39] A. Boccia, P. Falugi, H. Maurer, R.B. Vinter, Free time optimal control problems with time delays, In: Proceedings of the 52nd IEEE CDC, Florence, Italy, 2013, pp. 520-525.

[40] Q. Chai, R. Loxton, K.L. Teo, C. Yang, A unified parameter identification method for nonlinear time-delay systems, J. Ind. Manag. Optim., **9** (2013), 471–486.

[41] C. Liu, R. Loxton, K.L. Teo, Optimal parameter selection for nonlinear multistage systems with time-delays, Comput. Optim. Appl., **59** (2014), 285–306.

[42] 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.

[43] F.W. Gembicki, Performance and sensitivity optimization: A vector index approach, Ph.D. Thesis, Case Western Reserve University, Cleveland, OH, 1974.

[44] P.K. Shukla, On the normal boundary intersection method for generation of efficient front, In: Proceedings of 2007 ICCS, Part I, Lect. Notes Comput. Sci., **4487**, 2007, pp. 310–317.

[45] 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.

[46] C. Büskens, H. Maurer, SQP-methods for solving optimal control problems with control and state constraints: Adjoint variables, sensitivity analysis and real-time control, J. Comput. Appl. Math., **120** (2000), 85–108.

[47] Y. Zhao, M.A. Stadtherr, Rigorous global optimization for dynamic systems subject to inequality path constraints, Ind. Eng. Chem. Res., **50** (2011), 12678–12693.

[48] A. Bryson, Y.C. Ho, Applied Optimal Control, Halsted Press, New York, 1975.

[49] R.W.H. Sargent, G.R. Sullivan, The development of an efficient optimal control package, In: Proceedings of the 8th IFIP Conference on Optimization Techniques, Würzburg, Germany, 1977, pp. 158–168.

[50] C. Jiang, Q. Lin, C. Yu., K.L. Teo, G.R. Duan, An exact penalty method for free terminal time optimal control problem with continuous inequality constraints, J. Optim. Theory Appl., **154** (2012), 30–53.

[51] J. Nocedal, S.J. Wright, Numerical Optimization, Springer-Verlag, New York, 1999.

[52] Y. Mu, Z. Xiu, D. Zhang, A combined bioprocess of biodiesel production by lipase with microbial production of 1,3-propanediol by *Klebsiella pneumoniae*, Biochem. Eng. J., **40** (2008), 537–541.

[53] K. Schittkowski, NLPQLP: A Fortran implementation of a sequential quadratic programming algorithm with distributed and non-monotone line search — User guide, University of Bayreuth, Bayreuth, 2007.

[54] A.C. Hindmarsh, Large ordinary differential equation systems and software, IEEE Control Syst. Mag., **2** (1982), 24–30.

[55] J. Stoer, R. Bulirsch, Introduction to Numerical Analysis, Springer, New York, 1980.

[56] C. Liu, Z. Gong, K.L. Teo, E. Feng, Multi-objective optimization of nonlinear switched time-delay systems in fed-batch process, Appl. Math. Model., **40** (2016), 10533–10548.