Multi-objective optimization of nonlinear switched time-delay systems in fed-batch process

Chongyang Liu[a,b], Zhaohua Gong[a], Kok Lay Teo[b], Enmin Feng[c]

[a] *School of Mathematics and Information Science, Shandong Institute of Business and Technology, Yantai 264005, Shandong, China*
[b] *Department of Mathematics and Statistics, Curtin University, Perth 6845, Australia*
[c] *School of Mathematical Sciences, Dalian University of Technology, Dalian 116024, Liaoning, China*

Abstract

Maximization of productivity and minimization of consumption are two top priorities for biotechnological industry. In this paper, we model a fed-batch process as a nonlinear switched time-delay system. Taking the productivity of target product and the consumption rate of substrate as the objective functions, we present a multi-objective optimization problem involving the nonlinear switched time-delay system and subject to continuous state inequality constraints. To solve the multi-objective optimization problem, we first convert the problem into a sequence of single-objective optimization problems by using convex weighted sum and normal boundary intersection methods. A gradient-based single-objective solver incorporating constraint transcription technique is then developed to solve these single-objective optimization problems. Finally, a numerical example is provided to verify the effectiveness of the numerical solution approach. Numerical results show that the normal boundary intersection method in conjunction with the developed single-objective solver is more favourable than the convex weighted sum method.

Key words: Switched time-delay system; Multi-objective optimization; Convex weighted sum; Normal boundary intersection; Fed-batch process

1. Introduction

Fed-batch processes are extensively used in the biotechnological industry. To ensure product quality and economic viability of a fed-batch process, it is important to optimally design its operating strategy. Typically, in the literature, a single objective, e.g., maximization of product productivity, is considered [1, 2, 3, 4]. However, product productivity alone is not sufficient to provide a full indication of the economic viability of the fermentor, since the substrate consumption must also be taken into consideration. Thus, optimization of the

Email address: liu_chongyang@yahoo.com (Chongyang Liu)

fermentation processes is, in fact, a multi-objective optimization (MOO) problem.

1,3-Propanediol (1,3-PD) is a bulk chemical used as a monomer in the production of polyesters, polyethers and polyurethanes. These polymers possess fine qualities and will continue to be used widely in the future [5]. Production methods for 1,3-PD can be divided into two categories: chemical synthesis and microbial conversion. Compared with chemical synthesis, microbial conversion method has become increasingly attractive in the industry due to easy availability of renewable feedstock, such as glycerol – a byproduct of biodiesel production [6]. Glycerol is converted to 1,3-PD via bacterial fermentation [7]. Glycerol fermentation to produce 1,3-PD is a complex bioprocess, since the microbial growth is subjected to multiple inhibitions of substrate and products [5] and there exist time-delays in the process [8, 9]. Regarding the various fermentation techniques, which include batch mode, fed-batch mode and continuous mode, the fed-batch fermentation is typically implemented by switching between batch mode (in which the feeding of substrate is closed) and feeding mode (in which the feeding of substrate is open). This switching manner can reduce effectively the substrate inhibition and improve the 1,3-PD productivity, making the fed-batch fermentation being the most efficient cultivation method in 1,3-PD production [10]. The fed-batch process for converting glycerol to 1,3-PD begins with a batch operation. During this initial batch phase, the biomass tends to grow exponentially. Once the exponential growth phase ends, glycerol and alkali are added to the reactor to provide nutrition and regulate the pH level. The process then reverts to batch mode, and so on until the end of the final batch mode.

To achieve economically competitive production of 1,3-PD, optimization of the microbial conversion process is critical. Thus, many studies have been carried out on modelling and optimization of fed-batch process. The process is modelled as a nonlinear impulsive system in [11], where the addition of glycerol and alkali is assumed to be in an impulsive form. For this system, the corresponding parameter identification problem was investigated in [11, 12]. Furthermore, by taking the concentration of 1,3-PD at the terminal time as the objective function, a dynamic optimization problem was discussed in [13]. However, in reality, glycerol and alkali are added continuously. Thus, the fed-batch process is modelled as a nonlinear multistage system in [14]. Again, taking the concentration of 1,3-PD at the terminal time as the objective function, dynamic optimization problems involving the nonlinear multistage system were discussed in [14, 15, 16]. However, time-delays are ignored in the nonlinear systems mentioned above. In fact, like most real systems, fed-batch process is also influenced by time-delays. For this, a nonlinear multistage time-delay system was proposed in [17], where the corresponding parameter identification problem was also discussed. More recently, many important results obtained from the optimization of 1,3-PD production processes are summarized in [18]. However, only one objective function is involved in these optimization problems, meaning that they are single-objective optimization (SOO) problems.

In this paper, we model the 1,3-PD fed-batch process as a nonlinear switched

time-delay system with free terminal time. By taking both maximization of 1,3-PD productivity and minimization of consumption rate of substrate as the objectives, we then present a MOO problem involving this nonlinear switched time-delay system with free terminal time and subject to continuous state inequality constraints, where the feeding rate of glycerol, switching instants between batch and feeding modes, and duration of the fermentation process are regarded as the decision variables. As pointed out in [19, 20, 21], it is cumbersome to solve this free time optimization problem numerically, because numerical integration of the dynamic system must be conducted over a variable time interval at each optimization iteration. Thus, we introduce a time-scaling transformation [21] to convert the free time MOO problem into an equivalent one with fixed terminal time. However, unlike the case involving delay-free optimization problem, this time-scaling transformation causes the involving dynamic system to become a new switched system with variable time-delay. For the transformed MOO problem, it is conceptually different from a SOO problem. A key characteristic of a MOO problem is that the optimality is characterized by a set of solutions, called Pareto set, denoting the trade-offs between the competing objectives [22]. A solution is said to be Pareto optimal, if there is no other solution with better values of both objectives. Hence, when moving from one Pareto solution to another, any improvement in one objective can only occur with the worsening of at least one other. To generate the Pareto set of the MOO problem, two different approaches are often used [23]. The first approach, which is known as a scalarization method, transforms a MOO problem into a sequence of parametric SOO problems. By varying the parameters of the method involved, a representation of the Pareto set is obtained. This approach includes the classic convex weighted sum (CWS) [24] and normal boundary intersection (NBI) [25]. The second approach is referred to as a vectorization method. It utilizes heuristic optimization methods, such as genetic algorithm [26] and particle swarm optimization [27], to generate the Pareto set directly from the multi-objective formulation. Note that, for the scalarization approach, gradient-based deterministic optimization routes can be combined with to find (at least locally) optimal solutions for large-scale and highly constrained MOO problems in a fast and efficient way [28]. Consequently, such scalarization approach has been extensively used to solve MOO problems in biochemical processes [29, 30, 31, 32].

To solve the MOO problem considered in this paper, we transcribe the equivalent MOO problem into a sequence of SOO problems by using the CWS and the NBI methods. It should be noted that the existing single-objective solvers, including those solvers mentioned in [29, 30, 31, 32], only deal with problems involving ordinary differential systems and thus cannot be used to solve the resulting SOO problems, which involve switched time-delay systems. Furthermore, it is well known that variable switching times pose a significant challenge for conventional numerical optimization techniques [33, 34, 35]. To overcome this challenge, an extended version of the time-scaling transformation in [21] is typically used to map the variable switching times to fixed time points in a new time horizon. However, this extended time-scaling transformation is not applicable to switched time-delay systems [36]. For these reasons, we develop a new

single-objective solver to solve the resulting SOO problems. On the one hand, by employing the constraint transcription technique [37], we approximate the continuous state inequality constraints by constraints in canonical form. On the other hand, we derive the gradients of the objective functions and the constraint functions with respect to the decision variables. On this basis, the CWS and the NBI methods in conjunction with the new gradient-based single-objective solver are used to solve the MOO problem, respectively. A numerical example is used to verify the effectiveness of the solution approach. Numerical results show that the NBI method is more favourable than the CWS method.

The rest of the paper is organized as follows. Section 2 gives the switched time-delay system for describing the fed-batch process. Section 3 presents the MOO problem and its equivalent form. The numerical solution methods for the MOO problem are developed in Section 4. A numerical example is discussed in Section 5. Finally, Section 6 provides some concluding remarks.

2. Nonlinear switched time-delay system

The fed-batch process for converting glycerol to 1,3-PD includes two modes: batch mode and feeding mode. Based on the work in [17], the dynamic model for batch mode is given by

$$\begin{cases} \dot{x}_1(t) = \mu(x(t))x_1(t-\alpha), \\ \dot{x}_2(t) = -q_2(x(t))x_1(t-\alpha), \\ \dot{x}_3(t) = q_3(x(t))x_1(t-\alpha), \\ \dot{x}_4(t) = q_4(x(t))x_1(t-\alpha), \\ \dot{x}_5(t) = q_5(x(t))x_1(t-\alpha), \\ \dot{x}_6(t) = 0, \end{cases}$$

where t denotes the process time; $x(t) := (x_1(t), x_2(t), x_3(t), x_4(t), x_5(t), x_6(t))^\top$ is the state vector whose components 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 delay argument; $\mu(x(t))$ is the specific growth rate of cells; $q_2(x(t))$ is the specific consumption rate of substrate; and $q_\ell(x(t))$, $\ell = 3, 4, 5$, are, respectively, the specific formation rates of the reaction products 1,3-PD, acetic acid and ethanol. These functions are given as follows:

$$\begin{aligned} \mu(x(t)) &:= \frac{\Delta_1 x_2(t)}{x_2(t) + k_1} \prod_{\ell=2}^{5} \left(1 - \frac{x_\ell(t)}{x_\ell^*}\right), \\ q_\ell(x(t)) &:= m_\ell + Y_\ell \mu(x(t)) + \frac{\Delta_\ell x_2(t)}{x_2(t) + k_\ell}, \quad \ell = 2, 3, 4, \\ q_5(x(t)) &:= q_2(x(t)) \left(\frac{c_1}{c_2 + \mu(x(t))x_2(t)} + \frac{c_3}{c_4 + \mu(x(t))x_2(t)} \right), \end{aligned}$$

where Δ_1, k_1, m_ℓ, Y_ℓ, Δ_ℓ, k_ℓ, c_1, c_2, c_3, and c_4 are kinetic parameters and x_ℓ^* are the maximal residual concentrations of substrate and reaction products.

The dynamic model for feeding mode is given by

$$\begin{cases} \dot{x}_1(t) = \mu(x(t))x_1(t-\alpha) - D(x(t),\xi_j)x_1(t), \\ \dot{x}_2(t) = D(x(t),\xi_j)\Big(\dfrac{c_{s0}}{1+r} - x_2(t)\Big) - q_2(x(t))x_1(t-\alpha), \\ \dot{x}_3(t) = q_3(x(t))x_1(t-\alpha) - D(x(t),\xi_j)x_3(t), \\ \dot{x}_4(t) = q_4(x(t))x_1(t-\alpha) - D(x(t),\xi_j)x_4(t), \\ \dot{x}_5(t) = q_5(x(t))x_1(t-\alpha) - D(x(t),\xi_j)x_5(t), \\ \dot{x}_6(t) = (1+r)\xi_j, \end{cases}$$

where c_{s0} is the concentration of the initial feed of glycerol; r is the velocity ratio of adding alkali to glycerol; ξ_j is the feeding rate of glycerol in the jth feeding mode; and $D(x(t),\xi_j)$ is the dilution rate at time t defined as

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

The complete fed-batch process switches between batch and feeding modes, and starts and ends in batch mode. Let N be the number of feeding modes. Then, there are exactly $2N+1$ modes in total (N feeding modes, $N+1$ batch modes). Thus, the complete fed-batch process can be expressed as the following nonlinear switched time-delay system:

$$\begin{cases} \dot{x}(t) = f^i(x(t), x(t-\alpha), \xi), \quad t \in (\tau_{i-1}, \tau_i], \quad i = 1, \ldots, 2N+1, \\ x(t) = \phi(t), \quad t \le \tau_0, \end{cases} \tag{1}$$

where $x(t-\alpha) := (x_1(t-\alpha), x_2(t-\alpha), x_3(t-\alpha), x_4(t-\alpha), x_5(t-\alpha), x_6(t-\alpha))^\top$ is the delayed state; $\xi := (\xi_1, \xi_2, \ldots, \xi_N)^\top$ is the feeding rate vector of glycerol;

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

$$\begin{aligned} & f^{2j}(x(t), x(t-\alpha), \xi) \\ & := \begin{pmatrix} \mu(x(t))x_1(t-\alpha) - D(x(t),\xi_j)x_1(t) \\ D(x(t),\xi_j)\Big(\dfrac{c_{s0}}{1+r} - x_2(t)\Big) - q_2(x(t))x_1(t-\alpha) \\ q_3(x(t))x_1(t-\alpha) - D(x(t),\xi_j)x_3(t) \\ q_4(x(t))x_1(t-\alpha) - D(x(t),\xi_j)x_4(t) \\ q_5(x(t))x_1(t-\alpha) - D(x(t),\xi_j)x_5(t) \\ (1+r)\xi_j \end{pmatrix}, \quad j = 1, \ldots, N; \end{aligned}$$

$\tau_1, \ldots, \tau_{2N}$, are switching instants such that $0 = \tau_0 \leq \tau_1 \leq \cdots \leq \tau_{2N+1} = t_f$; t_f is the terminal time; and $\phi : R \to R^6$ is a given history function. Note that the state of system (1) does not undergo jumps at the switching instants.

The feeding rate of glycerol during feeding mode cannot be unbounded. Thus, define the set of admissible feeding rate vectors as

$$\Xi := \{\xi \in R^N |\ a_j \leq \xi_j \leq b_j,\ j = 1, \ldots, N\}, \tag{2}$$

where $a_j > 0$ and $b_j > 0$ are lower and upper bounds for the feeding rate of glycerol during the jth feeding mode. Any $\xi \in \Xi$ is called an admissible vector of feeding rates.

Since biological considerations limit the rate of switching, there are maximal and minimal time durations that are spent on each of the batch and feeding modes. On this basis, define the set of admissible switching instants and terminal times as

$$\begin{aligned}\Gamma := \{(\tau_1, \ldots, t_f)^\top \in R^{2N+1} |\ & \rho_i \leq \tau_i - \tau_{i-1} \leq \delta_i,\ i = 1, \ldots, 2N, \\ & \rho_{2N+1} \leq t_f \leq \delta_{2N+1}\},\end{aligned} \tag{3}$$

where ρ_i and δ_i are lower and upper bounds, respectively. Any $\tau := (\tau_1, \ldots, \tau_{2N+1})^\top \in \Gamma$ is called an admissible vector of switching instants and terminal time.

For system (1), there exists a unique continuous solution $x(\cdot|\xi, \tau)$ corresponding to each given pair $\xi \times \tau \in \Xi \times \Gamma$ on $(-\infty, +\infty)$ [38]. Moreover, the concentrations of biomass, glycerol, reaction products and volume of culture fluid must be restricted to biological meaningful ranges. Thus, define

$$x(t|\xi, \tau) \in W := \prod_{\ell=1}^{6} [x_{*\ell}, x_\ell^*], \quad t \in [0, t_f], \tag{4}$$

where $x_{*\ell}$ are, respectively, the lower thresholds for cell growth for biomass, glycerol, 1,3-PD, acetic acid, ethanol and volume of culture fluid, and x_ℓ^* are the corresponding upper thresholds (as used in the formula for $\mu(x(t))$).

3. Multi-objective optimization problem

The economic viability of any fermentation-based production process depends on the performance of the fermentor. There is no exception in the fed-batch process for converting glycerol to 1,3-PD. Thus, for the fed-batch process operated as described by system (1), we consider the following objective functions to characterize the fermentor performance:
(i) Maximizing 1,3-PD productivity:

$$\max\ \frac{x_3(t_f|\xi, \tau) x_6(t_f|\xi, \tau)}{t_f},$$

(ii) Minimizing consumption rate of substrate:

$$\min \ \frac{(x_6(t_f|\xi,\tau)-\phi_6(0))c_{s0}}{t_f}.$$

It should be noted that similar objectives have been considered in optimizing other biochemical processes, see for example [29, 30].

Let the two competing objectives be denoted as the following vector objective, which is to be minimized:

$$J(\xi,\tau) := \left(-\frac{x_3(t_f|\xi,\tau)x_6(t_f|\xi,\tau)}{t_f}, \frac{(x_6(t_f|\xi,\tau)-\phi_6(0))c_{s0}}{t_f} \right)^\top. \tag{5}$$

Furthermore, taking into consideration of constraints (2), (3) and (4), we present the following MOO problem:

$$\begin{array}{lll} \text{(MP)} & \min & J(\xi,\tau) \\ & \text{s.t.} & x(t|\xi,\tau) \in W, \ t \in [0,t_f], \\ & & (\xi,\tau) \in \Xi \times \Gamma, \end{array}$$

where W is as defined in (4). For (MP), there are three non-standard features: (i) the terminal time is free instead of fixed; (ii) the objective is not a scalar but a vector; and (iii) constraint (4) is a continuous state inequality constraint (i.e., it must be satisfied at an infinite number of points in the time interval).

To circumvent the first difficulty, we apply a time-scaling transformation [21] from $[0,t_f]$ to $[0,1]$ as given below:

$$t = t(s) = t_f s, \tag{6}$$

where $s \in [0,1]$ is a new time variable. Clearly, $s=0$ corresponds to $t=0$, $s=1$ corresponds to $t=t_f$ and the switching instants become $\theta_i := \tau_i/t_f, i = 1,\ldots,2N+1$. Under time-scaling transformation (6), system (1) is transformed into an equivalent form given below:

$$\begin{cases} \dot{\tilde{x}}(s) = \tilde{f}^i(\tilde{x}(s),\tilde{x}(s-t_f^{-1}\alpha),\xi,t_f), \ s \in (\theta_{i-1},\theta_i], \ i=1,\ldots,2N+1, \\ \tilde{x}(s) = \tilde{\phi}(s), \quad s \le \theta_0, \end{cases} \tag{7}$$

where $\tilde{x}(s) := x(t_f s)$; $\tilde{f}^i(\tilde{x}(s),\tilde{x}(s-t_f^{-1}\alpha),\xi,t_f) := t_f f^i(\tilde{x}(s),\tilde{x}(s-t_f^{-1}\alpha),\xi)$; and $\tilde{\phi}(s) := \phi(t_f s)$. Note that system (7) is of fixed terminal time and involves a variable time-delay. In addition, constraint (3) becomes

$$\Theta := \{(\theta_1,\ldots,\theta_{2N})^\top \in R^{2N} |\ \tilde{\rho}_i \le \theta_i - \theta_{i-1} \le \tilde{\delta}_i, \ i=1,\ldots,2N\}, \tag{8}$$

and

$$\mathcal{T} := \{t_f \in R |\ \rho_{2N+1} \le t_f \le \delta_{2N+1}\}, \tag{9}$$

where $\tilde{\rho}_i := \rho_i/t_f$; and $\tilde{\delta}_i := \delta_i/t_f$.

Let $\tilde{x}(\cdot|t_f,\xi,\theta)$ be the solution of system (7) corresponding to each given pair $(t_f,\xi,\theta)\in\mathcal{T}\times\Xi\times\Theta$ on $(-\infty,1]$. Then, constraint (4) turns into

$$\tilde{x}(s|t_f,\xi,\theta)\in W,\quad s\in[0,1], \tag{10}$$

and vector objective (5) becomes

$$\tilde{J}(t_f,\xi,\theta):=\left(-\frac{\tilde{x}_3(1|t_f,\xi,\theta)\tilde{x}_6(1|t_f,\xi,\theta)}{t_f},\frac{(\tilde{x}_6(1|t_f,\xi,\theta)-\tilde{\phi}_6(0))c_{s0}}{t_f}\right)^\top. \tag{11}$$

Therefore, (MP) can be stated equivalently as given below:

$$\begin{array}{lll}(\text{EMP}) & \min & \tilde{J}(t_f,\xi,\theta)\\ & \text{s.t.} & \tilde{x}(s|t_f,\xi,\theta)\in W,\ s\in[0,1],\\ & & (t_f,\xi,\theta)\in\mathcal{T}\times\Xi\times\Theta.\end{array}$$

Obviously, (EMP) is a MOO problem with fixed terminal time but involving a new switched system with variable time-delay.

4. Numerical solution approach

The optimal solutions of (EMP) are characterized by the Pareto set. In this section, we will develop numerical solution approach for (EMP).

4.1. Multi-objective optimization strategies

In this subsection, we shall introduce two strategies, i.e., CWS [24] and NBI [25], to convert (EMP) into a sequence of SOO problems.

4.1.1. Convex weighted sum

To generate Pareto set of a MOO problem, the most common scalarization approach is CWS. In this method, the individual objective functions are assigned different weights and then added together to form a single objective function. In essence, the objective weights provide additional degrees of freedom in the optimization problem. For (EMP), the corresponding SOO problem after the use of CWS is given by

$$\begin{array}{lll}(\text{EMP}_w) & \min & \tilde{J}_w(t_f,\xi,\theta):=\omega_1\tilde{J}_1(t_f,\xi,\theta)+\omega_2\tilde{J}_2(t_f,\xi,\theta)\\ & \text{s.t.} & \tilde{x}(s|t_f,\xi,\theta)\in W,\ s\in[0,1],\\ & & (t_f,\xi,\theta)\in\mathcal{T}\times\Xi\times\Theta,\end{array} \tag{12}$$

where $\omega:=(\omega_1,\omega_2)^\top$ is a vector of weights such that $\omega_1+\omega_2=1$ with $\omega_1,\omega_2\geq 0$.

4.1.2. Normal boundary intersection

NBI is a new scalarization scheme for generating Pareto set. It tackles MOO problem from a geometrically intuitive viewpoint. It first builds a plane in the objective space which contains all convex combinations of the individual minima, i.e., the convex hull of individual minima (CHIM), and then constructs quasi-normal lines to this plane. Thus, MOO problem is reformulated as to maximize the distance from a point on the CHIM along the quasi-normal through this point. Technically, this requirement of lying on the quasi-normal introduces additional equality constraints. By selecting the points on the CHIM, Pareto set in the objective space is obtained. As a result, (EMP) can be reformulated by using the NBI as

$$
\begin{aligned}
(\mathrm{EMP}_{\mathrm{NBI}}) \qquad \max \quad & \nu && (13)\\
\text{s.t.} \quad & \Phi\omega - \nu\Phi e = \tilde{J}(t_f,\xi,\theta) - \tilde{J}^*, && (14)\\
& \tilde{x}(s|t_f,\xi,\theta) \in W,\ s \in [0,1],\\
& (t_f,\xi,\theta) \in \mathcal{T}\times\Xi\times\Theta,
\end{aligned}
$$

where Φ is the pay-off matrix defined as

$$
\Phi := \begin{bmatrix} 0 & \tilde{J}_1(t_f^{2*},\xi^{2*},\theta^{2*}) - \tilde{J}_1(t_f^{1*},\xi^{1*},\theta^{1*}) \\ \tilde{J}_2(t_f^{1*},\xi^{1*},\theta^{1*}) - \tilde{J}_2(t_f^{2*},\xi^{2*},\theta^{2*}) & 0 \end{bmatrix},
$$

with $(t_f^{1*},\xi^{1*},\theta^{1*})$ and $(t_f^{2*},\xi^{2*},\theta^{2*})$ being the minimizers of objectives $\tilde{J}_1$ and $\tilde{J}_2$, respectively; ω is a vector of weights as used in (EMP_w); $e = (1,1)^\top$; and $\tilde{J}^* := (\tilde{J}_1(t_f^{1*},\xi^{1*},\theta^{1*}), \tilde{J}_2(t_f^{2*},\xi^{2*},\theta^{2*}))^\top$. Here, $\Phi\omega$ describes a point in the CHIM and $-\Phi e$ defines the quasi-normal to the CHIM towards $\tilde{J}^*$.

4.2. Single-objective solver

A sequence of resulting SOO problems must be solved when applying the CWS and the NBI strategies. In this subsection, we shall develop a gradient-based single-objective solver incorporating the constraint transcription technique [37] to solve these SOO problems.

4.2.1. Constraint transcription

As stated in Section 3, constraint (10) is a continuous state inequality constraint and this type of constraints is difficult to deal with in terms of numerical computation [17, 37]. We will apply a constraint transcription technique to transcribe constraint (10) into a canonical constraint.

To begin with, let

$$
\begin{aligned}
&h_\ell(\tilde{x}(s|t_f,\xi,\theta)) := x_\ell^* - \tilde{x}_\ell(s|t_f,\xi,\theta),\\
&h_{6+\ell}(\tilde{x}(s|t_f,\xi,\theta)) := \tilde{x}_\ell(s|t_f,\xi,\theta) - x_{*\ell},\ \ell = 1,\ldots,6.
\end{aligned}
$$

Then, constraint (10) is equivalent to the following equality constraint:

$$
\sum_{l=1}^{12}\int_0^1 \min\{0, h_l(\tilde{x}(s|t_f,\xi,\theta))\}ds = 0. \tag{15}
$$

However, since $\min\{0,\cdot\}$ is non-differentiable, constraint (15) is not suitable for gradient-based optimization techniques (which we will exploit later to design a solution algorithm). We therefore consider the following smooth approximation of the $\min\{0,\cdot\}$ function:

$$\min\{0,\eta\} \approx \pi_\epsilon(\eta),$$

where

$$\pi_\epsilon(\eta) = \begin{cases} \eta, & \text{if } \eta < -\epsilon, \\ -\dfrac{(\eta-\epsilon)^2}{4\epsilon}, & \text{if } -\epsilon \le \eta \le \epsilon, \\ 0, & \text{otherwise,} \end{cases}$$

and $\epsilon > 0$ is an adjustable parameter. It is easy to verify that π_ϵ is continuously differentiable and non-positive. Using π_ϵ, constraint (15) is approximated by

$$\tilde{G}^{\epsilon,\gamma}(t_f,\xi,\theta) := \gamma + \sum_{l=1}^{12} \int_0^1 \pi_\epsilon(h_l(\tilde{x}(s|t_f,\xi,\theta)))ds \ge 0, \tag{16}$$

where $\gamma > 0$ is an adjustable parameter. Hence, with this approximation scheme, constraint (10) is approximated by canonical constraint (16). This constraint is a standard constraint and can be readily handled using the gradient-based technique. In particular, by similar arguments as those given for Lemma 8.3.3 [37], it can be shown that for any $\epsilon > 0$, if $\gamma > 0$ is chosen sufficiently small, the solutions of the corresponding SOO problems with constraint (16) will also satisfy constraint (10).

4.2.2. Gradient formulas

To solve each of the resulting SOO problems by using gradient-based optimization techniques, the gradients of objective functions (12) and (13) as well as the gradients of constraint functions (8), (14) and (16) are needed. For the gradients of objective function (13) and constraint function (14) with respect to the decision variable ν, they can be easily calculated. The gradient of constraint function (8) with respect to θ is also obvious. Now, we note that objective (12) and constraint (14) are functions of $\tilde{J}$. Therefore, we only need to provide the gradients of $\tilde{J}$ and constraint function $\tilde{G}^{\epsilon,\gamma}$ with respect to t_f, ξ and θ.

First, define

$$\hat{f}^i(s|t_f,\xi,\theta) := \tilde{f}^i(\tilde{x}(s|t_f,\xi,\theta), \tilde{x}(s-t_f^{-1}\alpha|t_f,\xi,\theta),\xi,t_f),$$

$$\psi(s|t_f,\xi,\theta) := \begin{cases} t_f\dot{\phi}(t_f s), & s \le 0, \\ \hat{f}^i(s|t_f,\xi,\theta), & s \in (\theta_{i-1},\theta_i], \quad i = 1,\ldots,2N+1, \end{cases}$$

and

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

Then, the following theorem gives the gradient formulas of $\tilde{J}_1(t_f,\xi,\theta)$ with respect to t_f, ξ and θ.

Theorem 1. *Let* $(t_f, \xi, \theta) \in \mathcal{T} \times \Xi \times \Theta$. *Then*

$$\begin{aligned}\frac{\partial \tilde{J}_1(t_f,\xi,\theta)}{\partial t_f} &= \tilde{x}_3(1)\tilde{x}_6(1)t_f^{-2} + \sum_{i=1}^{2N+1}\int_{\theta_{i-1}}^{\theta_i} \lambda^\top(s) t_f^{-1}\hat{f}^i(s|t_f,\xi,\theta)ds \\ &+ \sum_{i=1}^{2N+1}\int_{\theta_{i-1}}^{\theta_i} \lambda^\top(s) t_f^{-2}\alpha \frac{\partial \hat{f}^i(s|t_f,\xi,\theta)}{\partial \tilde{x}(s-t_f^{-1}\alpha)}\psi(s-t_f^{-1}\alpha|t_f,\xi,\theta)ds \\ &+ \sum_{i=1}^{2N+1}\int_{\theta_{i-1}-t_f^{-1}\alpha}^{\theta_i-t_f^{-1}\alpha} \Big\{\lambda^\top(s+t_f^{-1}\alpha)s\frac{\partial \hat{f}(s+t_f^{-1}\alpha|t_f,\xi,\theta)}{\partial \tilde{x}(s)} \\ &\times \dot{\phi}(t_f s)\chi_{(-\infty,0)}(s)\Big\}ds, \qquad (17)\end{aligned}$$

$$\begin{aligned}\frac{\partial \tilde{J}_1(t_f,\xi,\theta)}{\partial \theta_i} &= \lambda^\top(\theta_i)\hat{f}^i(\theta_i|t_f,\xi,\theta) - \lambda^\top(\theta_i)\hat{f}^{i+1}(\theta_i|t_f,\xi,\theta), \\ &\qquad i = 1,\ldots,2N, \qquad (18)\end{aligned}$$

and

$$\frac{\partial \tilde{J}_1(t_f,\xi,\theta)}{\partial \xi_j} = \int_{\theta_{2j-1}}^{\theta_{2j}} \lambda^\top(s)\frac{\partial \hat{f}^{2j}(s|t_f,\xi,\theta)}{\partial \xi_j}ds, \quad j = 1,\ldots,N, \qquad (19)$$

where $\tilde{x}(1) := \tilde{x}(1|t_f,\xi,\theta)$*; and* $\lambda(\cdot) := \lambda(\cdot|t_f,\xi,\theta)$ *is the solution of the following costate system:*

$$\begin{aligned}\dot{\lambda}(s) &= -\sum_{i=1}^{2N+1}\left(\frac{\partial \hat{f}^i(s|t_f,\xi,\theta)}{\partial \tilde{x}(s)}\right)^\top \lambda(s)\chi_{(\theta_{i-1},\theta_i]}(s) \\ &- \sum_{i=1}^{2N+1}\left(\frac{\partial \hat{f}^i(s+t_f^{-1}\alpha|t_f,\xi,\theta)}{\partial \tilde{x}(s)}\right)^\top \lambda(s+t_f^{-1}\alpha)\chi_{(\theta_{i-1}-t_f^{-1}\alpha,\theta_i-t_f^{-1}\alpha]}(s), \\ &\qquad s \in [0,1], \qquad (20)\end{aligned}$$

with the terminal conditions

$$\lambda(1) = (0,0,-\tilde{x}_6(1)t_f^{-1},0,0,-\tilde{x}_3(1)t_f^{-1})^\top, \qquad (21)$$

$$\lambda(s) = (0,0,0,0,0,0)^\top, \quad s > 1. \qquad (22)$$

Proof. The derivations of the gradients of $\tilde{J}_1(t_f,\xi,\theta)$ with respect to θ_i, $i \in \{1,\ldots,2N\}$ and ξ_j, $j \in \{1,\ldots,N\}$ are similar to that of the gradient of $\tilde{J}_1(t_f,\xi,\theta)$ with respect to t_f. Thus, only the derivation of the gradient of $\tilde{J}_1(t_f,\xi,\theta)$ with respect t_f is given below.

Let $w : [0,\infty) \to R^6$ be an arbitrary function that is continuous and differ-

entiable almost everywhere. Then we may express $\tilde{J}_1(t_f, \xi, \theta)$ as follows:

$$\begin{aligned}
&\tilde{J}_1(t_f, \xi, \theta) \\
&= -t_f^{-1}\tilde{x}_3(1)\tilde{x}_6(1) + \sum_{i=1}^{2N+1} \int_{\theta_{i-1}}^{\theta_i} w(s)^\top \big(\hat{f}^i(s|t_f, \xi, \theta) - \dot{\tilde{x}}(s)\big) ds \\
&= -t_f^{-1}\tilde{x}_3(1)\tilde{x}_6(1) + \sum_{i=1}^{2N+1} \int_{\theta_{i-1}}^{\theta_i} w(s)^\top \hat{f}^i(s|t_f, \xi, \theta) ds - \int_0^1 w(s)^\top \dot{\tilde{x}}(s) ds,
\end{aligned}$$

where we have omitted the arguments t_f, ξ and θ in $\tilde{x}(\cdot|t_f, \xi, \theta)$ for brevity.

Applying integration by parts to the last integral term gives

$$\begin{aligned}
\tilde{J}_1(t_f, \xi, \theta) = -t_f^{-1}\tilde{x}_3(1)\tilde{x}_6(1) + \sum_{i=1}^{2N+1} \int_{\theta_{i-1}}^{\theta_i} w(s)^\top \hat{f}^i(s|t_f, \xi, \theta) ds - w(1)^\top \tilde{x}(1) \\
+ w(0)^\top \tilde{\phi}(0) + \int_0^1 \dot{w}(s)^\top \tilde{x}(s) ds. \qquad (23)
\end{aligned}$$

Differentiating (23) with respect to t_f gives

$$\begin{aligned}
&\frac{\partial \tilde{J}_1(t_f, \xi, \theta)}{\partial t_f} \\
&= t_f^{-2}\tilde{x}_3(1)\tilde{x}_6(1) - t_f^{-1}\tilde{x}_6(1)\frac{\partial \tilde{x}_3(1)}{\partial t_f} - t_f^{-1}\tilde{x}_3(1)\frac{\partial \tilde{x}_6(1)}{\partial t_f} - w(1)^\top \frac{\partial \tilde{x}(1)}{\partial t_f} \\
&+ \sum_{i=1}^{2N+1} \int_{\theta_{i-1}}^{\theta_i} w(s)^\top t_f^{-1} \hat{f}^i(s|t_f, \xi, \theta) ds + \sum_{i=1}^{2N+1} \int_{\theta_{i-1}}^{\theta_i} w(s)^\top \frac{\partial \hat{f}^i(s|t_f, \xi, \theta)}{\partial \tilde{x}(s)} \frac{\partial \tilde{x}(s)}{\partial t_f} ds \\
&+ \sum_{i=1}^{2N+1} \int_{\theta_{i-1}}^{\theta_i} w(s)^\top t_f^{-2}\alpha \frac{\partial \hat{f}^i(s|t_f, \xi, \theta)}{\partial \tilde{x}(s - t_f^{-1}\alpha)} \psi(s - t_f^{-1}\alpha|t_f, \xi, \theta) ds + \int_0^1 \dot{w}(s)^\top \frac{\partial \tilde{x}(s)}{\partial t_f} ds \\
&+ \sum_{i=1}^{2N+1} \int_{\theta_{i-1}}^{\theta_i} w(s)^\top \frac{\partial \hat{f}^i(s|t_f, \xi, \theta)}{\partial \tilde{x}(s - t_f^{-1}\alpha)} \frac{\partial \tilde{x}(s - t_f^{-1}\alpha)}{\partial t_f} ds. \qquad (24)
\end{aligned}$$

Performing a change of variable in the last term on the right-hand side of (24) yields

$$\begin{aligned}
&\sum_{i=1}^{2N+1} \int_{\theta_{i-1}}^{\theta_i} w(s)^\top \frac{\partial \hat{f}^i(s|t_f, \xi, \theta)}{\partial \tilde{x}(s - t_f^{-1}\alpha)} \frac{\partial \tilde{x}(s - t_f^{-1}\alpha)}{\partial t_f} ds \\
&= \sum_{i=1}^{2N+1} \int_{\theta_{i-1} - t_f^{-1}\alpha}^{\theta_i - t_f^{-1}\alpha} w(s + t_f^{-1}\alpha)^\top \frac{\partial \hat{f}^i(s + t_f^{-1}\alpha|t_f, \xi, \theta)}{\partial \tilde{x}(s)} \frac{\partial \tilde{x}(s)}{\partial t_f} ds. \qquad (25)
\end{aligned}$$

Since $\tilde{x}(s) = \tilde{\phi}(s)$ for all $s \leq 0$, (25) can be rewritten as follows:

$$\begin{aligned}
&\sum_{i=1}^{2N+1} \int_{\theta_{i-1}}^{\theta_i} w(s)^\top \frac{\partial \hat{f}^i(s|t_f,\xi,\theta)}{\partial \tilde{x}(s-t_f^{-1}\alpha)} \frac{\partial \tilde{x}(s-t_f^{-1}\alpha)}{\partial t_f} ds \\
&= \sum_{i=1}^{2N+1} \int_{\theta_{i-1}-t_f^{-1}\alpha}^{\theta_i-t_f^{-1}\alpha} w(s+t_f^{-1}\alpha)^\top \frac{\partial \hat{f}^i(s+t_f^{-1}\alpha|t_f,\xi,\theta)}{\partial \tilde{x}(s)} \chi_{[0,+\infty)}(s) \frac{\partial \tilde{x}(s)}{\partial t_f} ds \\
&+ \sum_{i=1}^{2N+1} \int_{\theta_{i-1}-t_f^{-1}\alpha}^{\theta_i-t_f^{-1}\alpha} w(s+t_f^{-1}\alpha)^\top s \frac{\partial \hat{f}^i(s+t_f^{-1}\alpha|t_f,\xi,\theta)}{\partial \tilde{x}(s)} \chi_{(-\infty,0)}(s) \dot{\phi}(t_f s) ds.
\end{aligned} \tag{26}$$

Substituting (26) into (24) yields

$$\begin{aligned}
&\frac{\partial \tilde{J}_1(t_f,\xi,\theta)}{\partial t_f} \\
&= t_f^{-2}\tilde{x}_3(1)\tilde{x}_6(1) - t_f^{-1}\tilde{x}_6(1)\frac{\partial \tilde{x}_3(1)}{\partial t_f} - t_f^{-1}\tilde{x}_3(1)\frac{\partial \tilde{x}_6(1)}{\partial t_f} - w(1)^\top \frac{\partial \tilde{x}(1)}{\partial t_f} \\
&+ \sum_{i=1}^{2N+1} \int_{\theta_{i-1}}^{\theta_i} w(s)^\top t_f^{-1} \hat{f}^i(s|t_f,\xi,\theta) ds + \sum_{i=1}^{2N+1} \int_{\theta_{i-1}}^{\theta_i} w(s)^\top \frac{\partial \hat{f}^i(s|t_f,\xi,\theta)}{\partial \tilde{x}(s)} \frac{\partial \tilde{x}(s)}{\partial t_f} ds \\
&+ \sum_{i=1}^{2N+1} \int_{\theta_{i-1}}^{\theta_i} w(s)^\top t_f^{-2}\alpha \frac{\partial \hat{f}^i(s|t_f,\xi,\theta)}{\partial \tilde{x}(s-t_f^{-1}\alpha)} \psi(s-t_f^{-1}\alpha|t_f,\xi,\theta) ds + \int_0^1 \dot{w}(s)^\top \frac{\partial \tilde{x}(s)}{\partial t_f} ds \\
&+ \sum_{i=1}^{2N+1} \int_{\theta_{i-1}-t_f^{-1}\alpha}^{\theta_i-t_f^{-1}\alpha} w(s+t_f^{-1}\alpha)^\top \frac{\partial \hat{f}^i(s+t_f^{-1}\alpha|t_f,\xi,\theta)}{\partial \tilde{x}(s)} \chi_{[0,+\infty)}(s) \frac{\partial \tilde{x}(s)}{\partial t_f} ds \\
&+ \sum_{i=1}^{2N+1} \int_{\theta_{i-1}-t_f^{-1}\alpha}^{\theta_i-t_f^{-1}\alpha} w(s+t_f^{-1}\alpha)^\top s \frac{\partial \hat{f}^i(s+t_f^{-1}\alpha|t_f,\xi,\theta)}{\partial \tilde{x}(s)} \chi_{(-\infty,0)}(s) \dot{\phi}(t_f s) ds.
\end{aligned}$$

Choosing $w(\cdot) = \lambda(\cdot|t_f,\xi,\theta)$ and substituting (20)-(22) into the above equation gives the gradient formula (17). The gradient formulas (18) and (19) can be derived similarly. The proof is completed. □

The next theorem gives the gradient formulas of $\tilde{J}_2(t_f,\xi,\theta)$ with respect to t_f, ξ and θ.

Theorem 2. *Let* $(t_f, \xi, \theta) \in \mathcal{T} \times \Xi \times \Theta$. *Then*

$$\begin{aligned}\frac{\partial \tilde{J}_2(t_f,\xi,\theta)}{\partial t_f} &= -(\tilde{x}_6(1) - \tilde{\phi}(0))c_{s0}t_f^{-2} + \sum_{i=1}^{2N+1}\int_{\theta_{i-1}}^{\theta_i} \bar{\lambda}^\top(s)t_f^{-1}\hat{f}^i(s|t_f,\xi,\theta)ds \\ &+ \sum_{i=1}^{2N+1}\int_{\theta_{i-1}}^{\theta_i} \bar{\lambda}^\top(s)t_f^{-2}\alpha\frac{\partial \hat{f}^i(s|t_f,\xi,\theta)}{\partial \tilde{x}(s-t_f^{-1}\alpha)}\psi(s-t_f^{-1}\alpha|t_f,\xi,\theta)ds \\ &+ \sum_{i=1}^{2N+1}\int_{\theta_{i-1}-t_f^{-1}\alpha}^{\theta_i-t_f^{-1}\alpha}\left\{\bar{\lambda}^\top(s+t_f^{-1}\alpha)s\frac{\partial \hat{f}^i(s+t_f^{-1}\alpha|t_f,\xi,\theta)}{\partial \tilde{x}(s)}\right. \\ &\left.\times\,\dot{\phi}(t_f s)\chi_{(-\infty,0]}(s)\right\}ds, \end{aligned} \tag{27}$$

$$\begin{aligned}\frac{\partial \tilde{J}_2(t_f,\xi,\theta)}{\partial \theta_i} &= \bar{\lambda}^\top(\theta_i)\hat{f}^i(\theta_i|t_f,\xi,\theta) - \bar{\lambda}^\top(\theta_i)\hat{f}^{i+1}(\theta_i|t_f,\xi,\theta), \\ &\qquad i = 1,\ldots,2N, \end{aligned} \tag{28}$$

and

$$\frac{\partial \tilde{J}_2(t_f,\xi,\theta)}{\partial \xi_j} = \int_{\theta_{2j-1}}^{\theta_{2j}} \bar{\lambda}^\top(s)\frac{\partial \hat{f}^{2j}(s|t_f,\xi,\theta)}{\partial \xi_j}ds, \quad j = 1,\ldots,N, \tag{29}$$

where $\tilde{x}(1) := \tilde{x}(1|t_f,\xi,\theta)$; *and* $\bar{\lambda}(\cdot) := \bar{\lambda}(\cdot|t_f,\xi,\theta)$ *is the solution of the following costate system:*

$$\begin{aligned}\dot{\bar{\lambda}}(s) &= -\sum_{i=1}^{2N+1}\left(\frac{\partial \hat{f}^i(s|t_f,\xi,\theta)}{\partial \tilde{x}(s)}\right)^\top \bar{\lambda}(s)\chi_{(\theta_{i-1},\theta_i]}(s) \\ &- \sum_{i=1}^{2N+1}\left(\frac{\partial \hat{f}^i(s+t_f^{-1}\alpha|t_f,\xi,\theta)}{\partial \tilde{x}(s)}\right)^\top \bar{\lambda}(s+t_f^{-1}\alpha)\chi_{(\theta_{i-1}-t_f^{-1}\alpha,\theta_i-t_f^{-1}\alpha]}(s), \\ &\qquad s \in [0,1], \end{aligned} \tag{30}$$

with the terminal conditions

$$\bar{\lambda}(1) = (0,0,0,0,0,c_{s0}t_f^{-1})^\top, \tag{31}$$

$$\bar{\lambda}(s) = (0,0,0,0,0,0)^\top, \quad s > 1. \tag{32}$$

Proof. The proof is similar to that given for Theorem 1 □.

The last theorem gives the gradient formulas of the constraint function $\tilde{G}^{\epsilon,\gamma}(t_f,\xi,\theta)$ with respect to t_f, ξ and θ.

Theorem 3. *Let $(t_f, \xi, \theta) \in \mathcal{T} \times \Xi \times \Theta$. Then for each $\epsilon > 0$ and $\gamma > 0$,*

$$\begin{aligned}\frac{\partial \tilde{G}^{\epsilon,\gamma}(t_f,\xi,\theta)}{\partial t_f} &= \sum_{i=1}^{2N+1} \int_{\theta_{i-1}}^{\theta_i} \tilde{\lambda}^\top(s) t_f^{-1} \hat{f}^i(s|t_f,\xi,\theta) ds \\ &+ \sum_{i=1}^{2N+1} \int_{\theta_{i-1}}^{\theta_i} \tilde{\lambda}^\top(s) t_f^{-2} \alpha \frac{\partial \hat{f}^i(s|t_f,\xi,\theta)}{\partial \tilde{x}(s - t_f^{-1}\alpha)} \psi(s - t_f^{-1}\alpha|t_f,\xi,\theta) ds \\ &+ \sum_{i=1}^{2N+1} \int_{\theta_{i-1}-t_f^{-1}\alpha}^{\theta_i - t_f^{-1}\alpha} \left\{ \tilde{\lambda}^\top(s + t_f^{-1}\alpha) s \frac{\partial \hat{f}^i(s + t_f^{-1}\alpha|t_f,\xi,\theta)}{\partial \tilde{x}(s)} \right. \\ &\left. \times \dot{\phi}(t_f s) \chi_{(-\infty,0]}(s) \right\} ds, \end{aligned} \tag{33}$$

$$\begin{aligned}\frac{\partial \tilde{G}^{\epsilon,\gamma}(t_f,\xi,\theta)}{\partial \theta_i} &= \tilde{\lambda}^\top(\theta_i) \hat{f}^i(\theta_i|t_f,\xi,\theta) - \tilde{\lambda}^\top(\theta_i) \hat{f}^{i+1}(\theta_i|t_f,\xi,\theta), \\ & \qquad i = 1, \ldots, 2N, \end{aligned} \tag{34}$$

and

$$\frac{\partial \tilde{G}^{\epsilon,\gamma}(t_f,\xi,\theta)}{\partial \xi_j} = \int_{\theta_{2j-1}}^{\theta_{2j}} \tilde{\lambda}^\top(s) \frac{\partial \hat{f}^{2j}(s|t_f,\xi,\theta)}{\partial \xi_j} ds, \quad j = 1, \ldots, N, \tag{35}$$

where $\tilde{\lambda}(\cdot) := \tilde{\lambda}(\cdot|t_f,\xi,\theta)$ is the solution of the following costate system:

$$\begin{aligned}\dot{\tilde{\lambda}}(s) &= -\sum_{l=1}^{12} \frac{\partial \pi_\epsilon(h_l(\tilde{x}(s|t_f,\xi,\theta)))}{\partial \tilde{x}(s)} - \sum_{i=1}^{2N+1} \left(\frac{\partial \hat{f}^i(s|t_f,\xi,\theta)}{\partial \tilde{x}(s)} \right)^\top \tilde{\lambda}(s) \chi_{(\theta_{i-1},\theta_i]}(s) \\ &- \sum_{i=1}^{2N+1} \left(\frac{\partial \hat{f}^i(s + t_f^{-1}\alpha|t_f,\xi,\theta)}{\partial \tilde{x}(s)} \right)^\top \tilde{\lambda}(s + t_f^{-1}\alpha) \chi_{(\theta_{i-1}-t_f^{-1}\alpha, \theta_i - t_f^{-1}\alpha]}(s), \\ & \qquad s \in [0,1], \end{aligned} \tag{36}$$

with the terminal condition

$$\tilde{\lambda}(s) = (0,0,0,0,0,0)^\top, \quad s \geq 1. \tag{37}$$

Proof. The proof is similar to that given for Theorem 1 □.

On the basis of Theorems 1-3, the gradients of objective function (12) and constraint functions (14) and (16) with respect to t_f, ξ and θ can be easily computed.

5. Numerical Example

Consider a 1,3-PD fed-batch production process by *Klebsiella pneumoniae* reported in [39]. This fed-batch process consists of an initial batch mode followed by some phases, where each phase involves an equal number of feeding and batch modes in succession. Within each phase, all feeding modes have the same

Table 1: Kinetic parameters, and maximal and minimal residual concentrations in system (1) [17].

ℓ	m_ℓ	Y_ℓ	Δ_ℓ	k_ℓ	c_ℓ	$x_{*\ell}$	x_ℓ^*
1	-	-	0.8	0.28	0.025	0.01	6
2	1.927	158.73	6.8489	17.7296	0.06	15	2039
3	-3.2819	80.6096	10.3687	15.50	2.81	0	1036
4	-0.97	33.07	5.74	85.71	65.5226	0	1026
5	-	-	-	-	-	0	360.9
6	-	-	-	-	-	4	6.55

duration and the same feeding rate of glycerol, and all batch modes, except the final one, have the duration of 100 seconds minus the feeding mode duration. Hence, the terminal time, the end moment of the first batch mode, the feeding rate of glycerol in each phase and the end moment of the first feeding mode in each phase are to be optimized.

To solve the corresponding SOO problems, we write a Fortran single-objective solver, where the gradients obtained in the previous section are incorporated in the optimization software NLPQLP – a Fortran implementation of sequential quadratic programming [40]. This single-objective solver invokes the differential equation software DLSODA [41] to solve the state and costate systems. Here, Lagrange interpolation [42] is used whenever DLSODA requires the value of the state or costate at an intermediate time between two adjacent knot points. The time-delay, the concentration of the initial feed of glycerol, the velocity ratio of adding alkali to glycerol, and the history function are $\alpha = 0.4652$h, $c_{s0} = 10672\text{mmolL}^{-1}$, $r = 0.75$, and $\phi(t) = (0.1115\text{gL}^{-1}, 495\text{mmolL}^{-1}, 0, 0, 0, 5\text{L})^\top$, respectively. The kinetic parameters and the maximal and minimal residual concentrations are listed in Table 1. Note that there are total 9 phases and 1335 switchings in the maximal duration of the fed-batch process. In 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 by τ_i, $i = 2, \ldots, 10$. The initial values and bounds of the switching instants and the terminal time are listed in Table 2, and the initial values and bounds of glycerol feeding rates are listed in Table 3. Moreover, the approximation parameters ϵ and γ are adjusted according to the ϵ–γ process as detailed in [21]. The process is terminated when $\epsilon \leq 1.0 \times 10^{-8}$ and $\gamma \leq 1.0 \times 10^{-7}$.

Both the CWS and the NBI methods are implemented with the single-objective solver described above using an even spread set of weights, in which the uniform spacing between two consecutive ω_1 is $1/19$ and $\omega_2 = 1 - \omega_1$ (resulting in 20 SOO problems), to solve (EMP). The obtained Pareto sets for these two methods are depicted in Figure 1. Note that a Pareto filter is designed to remove all non-global Pareto solutions whose definitions are given in [43]. This filter works by comparing a point in the Pareto set with every other generated

Table 2: Optimal values, initial values and lower and upper bounds for the switching instants and terminal time.

	Initial value	Lower bound	Upper bound	Weight $[1, 0]$	Weight $[0.684, 0.316]$	Weight $[0.316, 0.684]$	Weight $[0, 1]$
τ_1	5.3300	5.3000	5.4000	5.3020	5.3000	5.3000	5.3000
τ_2	6.1092	6.0783	6.1800	6.0804	6.0787	6.0792	6.0783
τ_3	7.1375	7.1067	7.2083	7.1087	7.1070	7.1076	7.1067
τ_4	8.8322	8.8011	8.9028	8.8040	8.8020	8.8026	8.8011
τ_5	12.137	-	12.208	12.110	12.108	12.107	12.098
τ_6	-	-	15.903	-	15.801	15.801	15.081
τ_7	-	-	18.152	-	-	18.050	18.042
τ_8	-	-	19.902	-	-	19.800	19.750
τ_9	-	-	23.902	-	-	23.800	23.780
τ_{10}	-	-	24.106	-	-	24.105	24.095
t_f	14.160	11.000	24.160	14.090	16.160	24.160	24.149

Table 3: Bounds of glycerol feeding rates in phases 1-9 [18].

Phases	1-2	3	4-5	6	7	8-9
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

points. If a point is not a global Pareto solution, then it is eliminated. From Figure 1, we can see that minimizing the minus productivity of 1,3-PD and minimizing the consumption rate of substrate are conflicting objectives because lowering one results in an increase of the other. Although both the CWS and the NBI methods capture this feature, there are large differences in the accuracy of the resulting Pareto sets. Although an even spread is used to vary the weights in the CWS and the NBI methods, the result obtained by the NBI method clearly exhibits a more even spread in the Pareto set than the one obtained by the CWS method. Thus, the result obtained by the NBI method provides a more accurate representation of Pareto set. In addition, it is observed that the computation time taken to generate one point by the NBI method is about 31.3s. However, it takes 50.8s by the CWS method. These computations are carried out on a computer equipped with an Intel Core i5-2300 CPU (2.80GHz) and 4.00GB RAM. Thus, it is clear that the NBI method not only generates a more accurate Pareto set than the CWS method, but also gives rise to faster computation time.

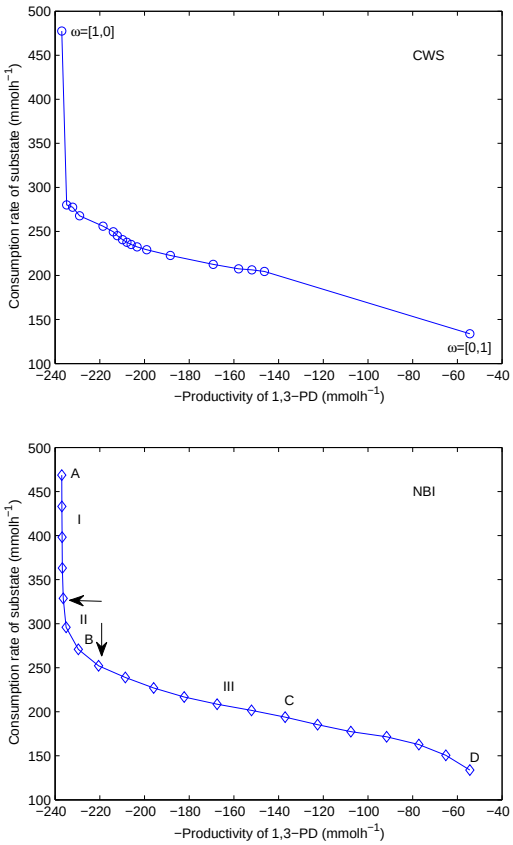
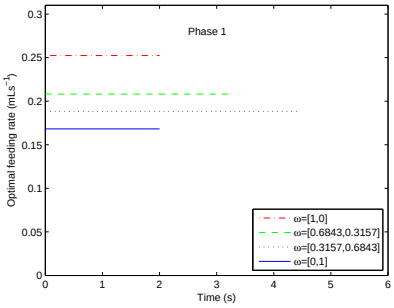
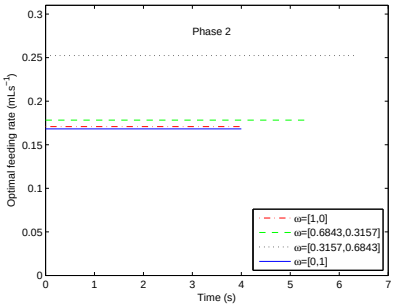
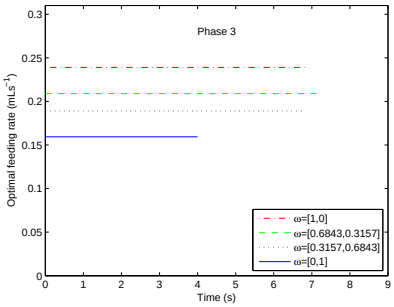


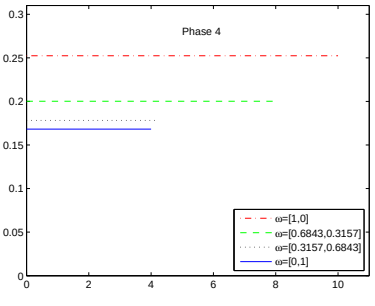
Figure 1: Pareto sets generated by CWS and NNC methods.

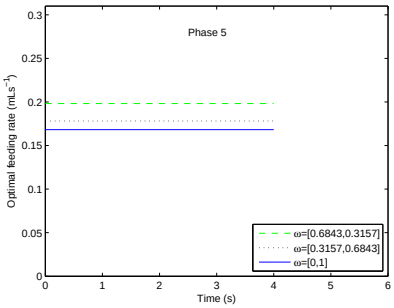
It should also be noted that the Pareto set obtained by the NBI method consists of three regions (see Figure 1). Region I is characterized by the limited ability of the consumption rate of substrate to 1,3-PD productivity, that is, a large increase in the consumption rate of substrate results in an insignificant increase in 1,3-PD productivity. For example, an increase in the consumption rate of substrate from 328.746mmolh^{-1} to 468.641mmolh^{-1} only gives rise to an increase of 1,3-PD productivity from 236.349mmolh^{-1} to 237.018mmolh^{-1}. Region III is characterized by the sensitivity of 1,3-PD productivity to the consumption rate of substrate, i.e., a small increase in the consumption rate of substrate produces a substantial increase in 1,3-PD productivity. For example, an increase in the consumption rate of substrate from 133.851mmolh^{-1} to 252.163mmolh^{-1} produces an increase in the 1,3-PD productivity from 54.412mmolh^{-1} to 220.547 mmolh^{-1}. Region II represents a transition zone between regions I and III. Furthermore, four points A, B, C, and D are taken from the Pareto set obtained

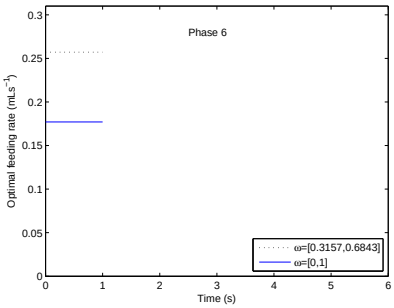

Phase 1
Optimal feeding rate (mLs^{-1})
Time (s)
ω=[1,0]
ω=[0.6843,0.3157]
ω=[0.3157,0.6843]
ω=[0,1]


Phase 2
Optimal feeding rate (mLs^{-1})
Time (s)
ω=[1,0]
ω=[0.6843,0.3157]
ω=[0.3157,0.6843]
ω=[0,1]


Phase 3
Optimal feeding rate (mLs^{-1})
Time (s)
ω=[1,0]
ω=[0.6843,0.3157]
ω=[0.3157,0.6843]
ω=[0,1]


Phase 4
ω=[1,0]
ω=[0.6843,0.3157]
ω=[0.3157,0.6843]
ω=[0,1]


Phase 5
Optimal feeding rate (mLs^{-1})
Time (s)
ω=[0.6843,0.3157]
ω=[0.3157,0.6843]
ω=[0,1]


Phase 6
Optimal feeding rate (mLs^{-1})
Time (s)
ω=[0.3157,0.6843]
ω=[0,1]

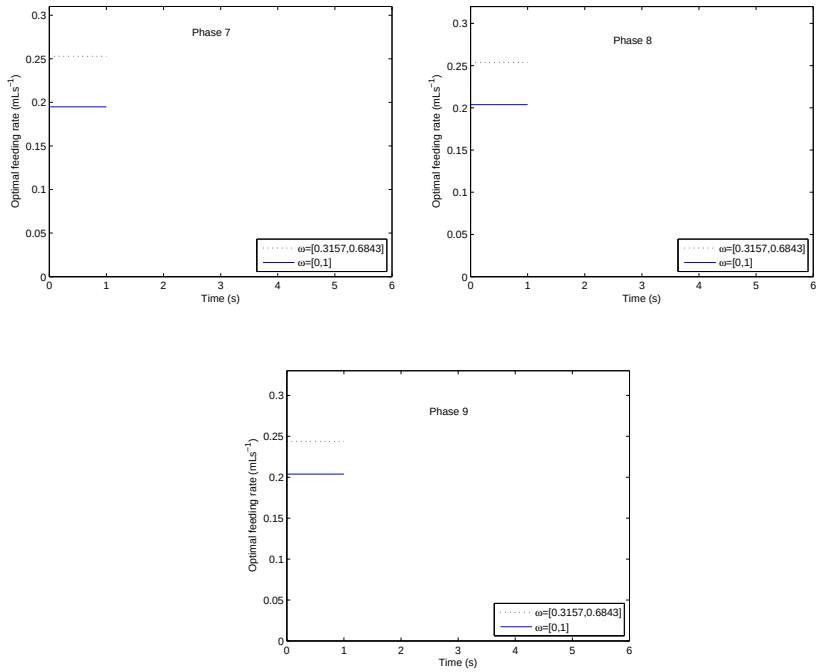


Figure 2: Optimal feeding rates of glycerol in phases 1-9 for points A, B, C and D.

by the NBI method (see Figure 1). For these four points, the corresponding weights, the optimal switching instants and terminal times are also listed in Table 2, and the optimal feeding rates of glycerol and the corresponding feeding durations in phases 1-9 are depicted in Figure 2. Note that the feeding durations of glycerol are in seconds and the batch durations are 100 seconds minus the corresponding feeding duration. Under these optimal profiles, the changes of 1,3-PD productivity and the consumption rate of substrate are depicted in Figure 3. From Figure 3, we can see that there is a trade-off between 1,3-PD productivity and the consumption rate of substrate.

6. Conclusions

This paper has studied MOO problems arising in 1,3-PD fed-batch process. Taking 1,3-PD productivity and the consumption rate of substrate as the objective functions, we presented a MOO optimization problem involving a switched time-delay system. To solve the MOO problem, we converted the problem into a sequence of SOO problems by using the CWS and the NBI methods. A single-objective solver was designed to solve the resulting SOO problems. A numerical example was used to test the developed numerical solution approach.

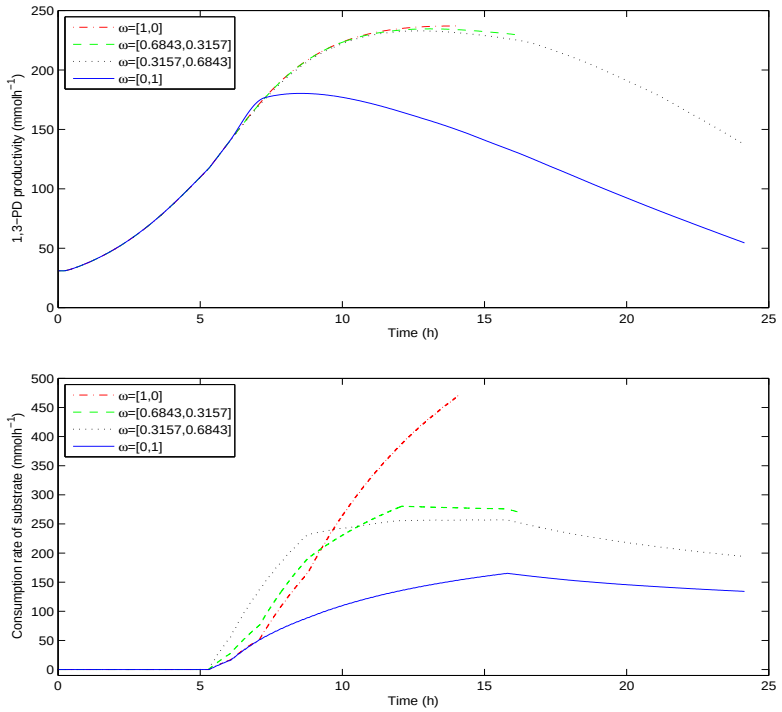


Figure 3: Changes of productivity of 1,3-PD and consumption rate of substrate with respect to fermentation time for points A, B, C and D.

Numerical results show that the NBI method provides a more accurate Pareto set than the CWS method.

Acknowledgments

This work is supported by the Natural Science Foundation of China (Nos. 11201267, 1117050), the Australian Research Council (No. DP140100289) and the Shandong Province Natural Science Foundation of China (Nos. ZR2015AL010, ZR2015PG006)

References

[1] A. Johnson, The control of fed-batch fermentation processes – A survey, Automatica 23 (1987) 691–705.

[2] R. Luus, D. Hennessy, Optimization of fed-batch reactors by the Luus-Jaakola optimization procedure, Ind. Eng. Chem. Res. 38 (1999) 1948–1955.

[3] D. Sarkar, J.M. Modak, Genetic algorithms with filters for optimal control problems in fed-batch bioreactors, Bioproc. Biosyst. Eng. 26 (2004) 295–306.

[4] E. Franco-Lara, D. Weuster-Botz, Estimation of optimal feeding strategies for fed-batch bioprocesses, Bioproc. Biosyst. Eng. 27 (2005) 255–262.

[5] 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.

[6] R. Saxena, P. Anand, S. Saran, J. Isar, Microbial production of 1,3-propanediol: Recent developments and emerging opportunities, Biotechnol. Adv. 27 (2009) 895–913.

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

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

[9] 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.

[10] A. Zeng, W. Sabra, Microbial production of diols as platform chemicals: Recent progresses, Curr. Opin. Biotech. 22 (2011) 749–757.

[11] C. Gao, E. Feng, Z. Wang, Z. Xiu, Nonlinear dynamical systems of biodissimilation of glycerol to 1,3-propanediol and their optimal controls, J. Ind. Manage. Optim. 1 (2005) 377–388.

[12] G. Wang, E. Feng, Z. Xiu, Modelling and parameter identification of microbial bioconversion in fed-batch cultures, J. Process Control 18 (2008) 458–464.

[13] 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.

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

[15] Z. Gong, C. Liu, E. Feng, L. Wang, Y. Yu, Modelling and optimization for a switched system in microbial fed-batch culture, Appl. Math. Model. 35 (2011) 3276–3284.

[16] C. Liu, Z. Gong, B. Shen, E. Feng, Modelling and optimal control for a fed-batch fermentation process, Appl. Math. Model 37 (2013) 695–706.

[17] 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.

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

[19] 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.

[20] 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.

[21] 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.

[22] K. Miettinen, Nonlinear Multiobjective Optimization, Kluwer Academic Publisher, Boston, 1999.

[23] R.T. Marler, J.S. Arora, Survey of multi-objective optimization methods for engineering, Struct. Multidiscip. Optim. 26 (2004) 369–395.

[24] L. Zadeh,: Optimality and non-scalar-valued performance criteria, IEEE Trans. Autom. Control 8 (1963) 59–60.

[25] 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.

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

[27] 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.

[28] F. Logist, B. Houska, M. Diehl, J. Van Impe, Fast Pareto set generation for nonlinear optimal control problems with multiple objectives, Struct. Multidiscip. Optim. 42 (2010) 591–603.

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

[30] 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.

[31] F. Logist, B. Houska, M. Diehl, J. Van Impe, A toolkit for efficiently generating Pareto sets in (bio)chemical multi-objective optimal control problems, Comput. Aided Chem. Eng. 28 (2010) 481–486.

[32] 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.

[33] R. Li, K.L. Teo, K.H. Wong, G.R. Duan, Control parameterization enhancing transform for optimal control of switched systems, Math. Comput. Modelling 43 (2006) 1393–1403.

[34] R. Loxton, K.L. Teo, V. Rehbock, W.K. Ling, Optimal switching instants for a switched-capacitor DC/DC power converter, Automatica 45 (2009) 973–980.

[35] 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.

[36] 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.

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

[38] J.K. Hale, S.M. Verduyn Lune, Introduction to Functional-Differential Equations, Springer, Berlin, 1993.

[39] 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.

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

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

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

[43] 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.