

An input/output approach to the optimal transition control of a class of distributed chemical reactors

Mingheng Li, Panagiotis D. Christofides*

Department of Chemical and Biomolecular Engineering, University of California, Los Angeles, CA 90095-1592, USA

Received 20 July 2006; received in revised form 12 February 2007; accepted 28 February 2007

Available online 14 March 2007

Abstract

An input/output approach to the optimal concentration transition control problem of a certain type of distributed chemical reactors is proposed based on the concept of residence time distribution, which can be determined in practice by using data from experimental measurements or computer simulations. The main assumptions for the proposed control method to apply are that the thermal and fluid flow fields in the reactor are at pseudo-steady-state during transition and that the component whose concentration is to be controlled participates only in first-order reactions. Using the concept of cumulative residence time distribution, the output variable is expressed as the weighted sum of discretized inputs or input gradients in order to construct an input/output model, on the basis of which a constrained optimal control problem, penalizing a quadratic control energy functional in the presence of input constraints, is formulated and solved as a standard least squares problem with inequality constraints. The effectiveness of the proposed optimal control scheme is demonstrated through a continuous-stirred-tank-reactor (CSTR) network and a tubular reactor with axial dispersion and a first-order reaction. It is demonstrated through computer simulations that the proposed control method is advantageous over linear quadratic regulator (LQR) and proportional-integral (PI) control in terms of control cost minimization and input constraint satisfaction.

© 2007 Elsevier Ltd. All rights reserved.

Keywords: Optimal control; Computational fluid dynamics; Residence time distribution; Chemical reactor; Concentration transition

1. Introduction

Modern chemical plants generally produce a variety of products, and it is common that these products are only slightly different in terms of their composition. For example, polymers of different molecular weights in a polyethylene plant have different properties such as texture and mechanical strength (e.g., Cervantes et al., 2002; Lo and Ray, 2006; McAuley and MacGregor, 1992). Glasses with different levels of colorant agent exhibit different appearances and solar performances (Trier, 1987). To satisfy the need of different customers, a modern manufacturing plant should be able to make product transitions in a short time and to maintain a high quality of the new product at the same time. In some cases, a product transition in a

chemical plant might take days or even weeks due to a long residence time of the chemical reactor involved. This makes the optimal control of outlet concentration in distributed chemical reactors an important subject at the interface of reactor engineering and process control. Due to the coupled flow, thermal and reaction phenomena involved in a distributed chemical reactor, a detailed description of flow, temperature and concentration fields in typical reactors require multi-dimensional computational fluid dynamic (CFD) simulations (e.g., Li and Christofides, 2005, 2006; Li et al., 2006). A direct use of CFD models for control design or dynamic optimization involves significant computational cost. Although application of advanced model reduction techniques to derive reduced-order models from the detailed partial differential equation (PDE) process models may work well in certain cases and lead to efficient dynamic optimization and control algorithms (see, for example, Armaou and Christofides, 1999, 2000, 2002; Baker et al., 2000; Baker and Christofides, 2000; Bendersky and Christofides, 2000; Christofides, 2001; Christofides and

* Corresponding author. Tel.: +1 310 794 1015; fax: +1 310 206 4107.

E-mail addresses: limh@ucla.edu (M. Li),
pdc@seas.ucla.edu (P.D. Christofides).

Daoutidis, 1997; Graham and Kevrekidis, 1996; Graham et al., 1999; Groetsch et al., 2006; Park and Lee, 2000), such a reduced-order model approach might require a huge amount of memory and computational cost when the CFD model consists of millions of grid points needed to accurately describe the process behavior. The reader may also refer to Raja et al. (2000) and Varshney and Armaou (2006) for recent applications of model reduction and dynamic optimization to thin film deposition processes described by CFD equations and Balsa-Canto et al. (2004, 2005) for further recent results on dynamic optimization and control of distributed parameter systems.

To circumvent the computational complexity brought by the possible high dimensionality of the dynamic system, an input/output approach to transition control problems for a class of distributed chemical reactors is proposed based on the concept of residence time distribution. Specifically, we consider processes in which the grade of the final product is determined by the concentration of a key component that has a very low concentration and participates only in first-order reactions. When these conditions are satisfied, the dynamics of the transition process is linear and an input/output model that describes the dynamic relationship between the inlet and outlet concentrations of the key component can be constructed using the concept of cumulative residence time distribution. A major advantage of this method is that the residence time distribution is a common term in reactor engineering and reliable measurement techniques are readily available (Fogler, 1998). The reader may also refer to several papers in the literature regarding the determination of the residence time distribution through CFD simulations (see Ekambara and Joshi, 2004; Gunjal et al., 2003; Harris et al., 1996; Heibel et al., 2001, 2005; Rigopoulos and Jones, 2003). Another advantage of the input/output approach is that once the residence time distribution is determined and the input/output model is constructed, iteratively solving the process dynamics (which may be described by PDEs) is not necessary. Finally, constraints on the input can be easily incorporated in the input/output model and handled in the solution of the transition control problem. Using the concept of cumulative residence time distribution, the output variable is expressed as the weighted sum of discretized inputs or input gradients in order to construct an input/output model, on the basis of which a constrained optimal control problem, penalizing a quadratic control energy functional in the presence of input constraints, is formulated and solved as a standard least squares problem with inequality constraints. The effectiveness of the proposed optimal control scheme is demonstrated through a continuous-stirred-tank-reactor (CSTR) network and a tubular reactor with axial dispersion and a first-order reaction. It is demonstrated that the proposed control method yields the same control trajectory as the one obtained from linear quadratic regulator (LQR) theory (designed based on the state-space model) if there are no constraints present. However, when constraints are present, the proposed optimal control method is shown to be advantageous over LQR and proportional-integral (PI) control in terms of control cost minimization and input constraint satisfaction.

2. Control problem formulation

The focus of this work is on a certain type of optimal transition control problems in a class of distributed chemical reactors where the grade of the product is directly related to the concentration of a key component in the effluent stream. This key component is assumed to participate only in first-order reactions although there might be other reactions (which can be of any order) occurring within the reactor. In such a case, the concentration in the exit of the reactor is regulated by the concentration at the entrance of the reactor. In general, the reactor might be non-isothermal and flow recirculation might exist within the reactor. However, a pseudo-steady-state of the flow and thermal fields should hold during the transition process. This assumption is valid if the concentration of the key component to be controlled is very low. It is also assumed that the concentration distribution at the plane of the exit is negligible. Considering such a reactive flow process with species transport, an accurate description of the transport phenomena involves the following dynamic conservation equations (the number of species is assumed to be two for simplicity):

$$\begin{aligned} \frac{\partial(\rho v)}{\partial t} + \nabla \cdot (\rho v v) &= \nabla \cdot (\boldsymbol{\tau}) - \nabla p + \mathbf{F}, \\ \frac{\partial(\rho h)}{\partial t} + \nabla \cdot (\rho h v) &= \nabla \cdot (\lambda/c_p \nabla h) S, \\ \frac{\partial(\rho \alpha_i)}{\partial t} + \nabla \cdot (\rho \alpha_i v) &= -\nabla \cdot (j_i) + R_i, \quad i = 1, 2, \end{aligned} \quad (1)$$

where ρ is the density of the fluid, $\mathbf{v} = [v_x, v_y, v_z]^T$ is the velocity vector, $\boldsymbol{\tau}$ is the stress, p is the pressure, h is the enthalpy, λ is the thermal conductivity, c_p is the heat capacity, α_i is fraction of component i in the mixture, and j_i is the diffusion flux of component i . \mathbf{F} , S , and R_i are the source terms. The manipulated input and controlled output variables appear in the boundary conditions of the above PDEs (see the second example in the case studies section for a detailed formulation). Through discretization in space (usually with local discretization techniques such as finite difference, finite element or finite volume, etc.) and incorporation of the boundary conditions, the above PDEs can be converted into a set of ordinary differential equations of the following form:

$$\begin{aligned} \dot{\mathbf{x}} &= \mathbf{f}(\mathbf{x}) + \mathbf{g}(\mathbf{x})u, \\ y &= \mathbf{h}(\mathbf{x}), \end{aligned} \quad (2)$$

where $\mathbf{x} = [\rho v_x^1, \rho v_y^1, \rho v_z^1, \rho h^1, \rho \alpha_1^1, \rho \alpha_2^1, \dots, \rho v_x^N, \rho v_y^N, \rho v_z^N, \rho h^N, \rho \alpha_1^N, \rho \alpha_2^N]^T$ is a $6N$ vector. Here N is the number of grid points in the CFD model. N might be several million in some large-scale chemical reactors. u and y are the concentration of the key component at the entrance and the exit of the reactor, respectively. The control problem is to minimize the following functional subject to the process dynamics and constraints

in the input:

$$\begin{aligned} \min_{u(t)} \quad & J = \int_0^\infty (y(t) - y_f)^2 dt + \varepsilon^2 \int_0^\infty (u(t) - u_f)^2 dt \\ \text{s.t.} \quad & \dot{\mathbf{x}}(t) = f(\mathbf{x}(t)) + g(\mathbf{x}(t))u, \\ & y(t) = h(\mathbf{x}(t)), \\ & u_{\min} \leq u(t) \leq u_{\max}, \end{aligned} \quad (3)$$

where u_f and y_f are the steady-state concentration of the key component at the inlet and outlet of the reactor after transition, and ε represents the weight on the control action during the transition process. Apparently, a steady-state solution of the process requires solving the full $6N$ dimensional state-space model. Because a pseudo-steady-state of the thermal and flow fields can be assumed during the transition process, the velocity and temperature do not need to be solved again in the dynamic simulation once they become available in the steady-state simulation. In such a case, a description of the concentration evolution within the reactor consists of, however, still N state variables. If N exceeds 10^6 , a huge amount of memory and computational cost is required in the control design although the dynamic model is linear.

3. Input/output approach to the control problem

To circumvent the computational complexity brought by the possible high dimensionality of the control problem described by Eq. (3) in the control design, it is proposed in this work to develop a model describing the input/output behavior through the concept of residence time distribution. Under the assumptions made in this work, the transition process is linear and the residence time distribution acts as an impulse response of a linear dynamic system. In this section, the transition process without reaction is first discussed, followed by one with a first-order reaction.

Let $p(t)$ be the residence time distribution in a generic reactor. The amount of the molecules that enter the reactor at time θ and spend time $t - \theta$ in the reactor before exiting is given by $u(\theta)p(t - \theta) d\theta$ if no reaction is involved. Therefore, the effluent concentration $y(t)$ is the convolution of $u(\theta)$ and $p(\theta)$, or

$$\begin{aligned} y(t) &= \int_{-\infty}^t u(\theta)p(t - \theta) d\theta \\ &= \int_0^\infty u(t - s)p(s) ds, \end{aligned} \quad (4)$$

where $s = t - \theta$. The cumulative residence time distribution $P(t)$ is related to $p(t)$ by

$$P(t) = \int_0^t p(\theta) d\theta. \quad (5)$$

For a reactor with no dead zones or sinks, the molecules entering the reactor will eventually exit from the effluent stream, i.e., $\lim_{t \rightarrow \infty} P(t) = 1$ and $\lim_{t \rightarrow \infty} p(t) = 0$. Let \mathcal{L} be the operator mapping $u(t)$ to $y(t)$, or $y(t) = \mathcal{L}([u(t)])$, it is readily

verified that

$$\begin{aligned} \mathcal{L}(a[u(t)]) &= a\mathcal{L}([u(t)]), \\ \mathcal{L}(a_1[u_1(t)] + a_2[u_2(t)]) &= a_1\mathcal{L}([u_1(t)]) + a_2\mathcal{L}([u_2(t)]), \\ \lim_{t \rightarrow \infty} \mathcal{L}(1) &= 1, \end{aligned} \quad (6)$$

where a , a_1 , and a_2 are constants and $u(t)$, $u_1(t)$, and $u_2(t)$ are the input variables.

Let y_0 and y_f be the steady-state value of the initial and final concentration of the key component in the exhaust during the transition process, and introduce two dimensionless variables $\bar{u}(t)$ and $\bar{y}(t)$ of the following form:

$$\begin{aligned} \bar{u}(t) &= \frac{u(t) - y_0}{u_f - u_0}, \\ \bar{y}(t) &= \frac{y(t) - y_0}{u_f - u_0}; \end{aligned} \quad (7)$$

it can be verified that

$$\bar{y}(t) = \mathcal{L}([\bar{u}(t)]), \quad (8)$$

provided that the process before transition is at steady-state, which guarantees $\mathcal{L}(1) = 1$. The constraints applied on $u(t)$ can be determined using the relationship $u(t) = y_0 + (u_f - u_0)\bar{u}(t)$. For example, if

$$u_{\min} \leq u(t) \leq u_{\max} \quad (9)$$

then

$$\begin{cases} \frac{u_{\min} - y_0}{u_f - u_0} \leq \bar{u}(t) \leq \frac{u_{\max} - y_0}{u_f - u_0} & (\text{if } y_0 < y_f), \\ \frac{u_{\max} - y_0}{u_f - u_0} \leq \bar{u}(t) \leq \frac{u_{\min} - y_0}{u_f - u_0} & (\text{if } y_f < y_0), \end{cases} \quad (10)$$

which can also be written in the following form:

$$\bar{u}_{\min} \leq \bar{u}(t) \leq \bar{u}_{\max}. \quad (11)$$

From the above analysis, it is seen that the concentration transition problem can be converted into a dimensionless form with the output changing from 0 to 1. In the remainder, we focus on the dimensionless form of this dynamic process and drop the bar sign on the input and output variables. The optimal control problem using the input/output model can be then expressed as follows:

$$\begin{aligned} \min_{u(t)} \quad & J = \int_0^\infty (y(t) - y_f)^2 dt + \varepsilon^2 \int_0^\infty (u(t) - u_f)^2 dt \\ \text{s.t.} \quad & y(t) = \int_{-\infty}^t u(\theta)p(t - \theta) d\theta, \\ & u_{\min} \leq u(t) \leq u_{\max}, \end{aligned} \quad (12)$$

where $y_f = u_f = 1$ is the target concentration of the input and output at the steady-state. The problem of Eq. (12) approximates well the original one of Eq. (3), which is based on a state-space formulation, provided that the residence time distribution does not change during the transition process.

For the processes under consideration, this condition is a reasonable one since the existence of the key component has negligible effect on the thermal and fluid fields in the reactor.

Solving Eq. (12) might involve a significant computational cost if the numerical discretization is not appropriately chosen. In this work, a method is proposed to represent the output as a linear sum of the inputs or the gradient of inputs, which is demonstrated to be computationally efficient in the example section. Specifically, if the input is discretized into n uniformly distributed time intervals, the output at time $t_{n-1} < t \leq t_n$ can be solved in the following way:

$$\begin{aligned}
 y(t) &= \sum_{i=1}^{n-1} \int_{t_{i-1}}^{t_i} u(\theta) p(t - \theta) d\theta + \int_{t_{n-1}}^t u(\theta) p(t - \theta) d\theta \\
 &= \sum_{i=1}^{n-1} u_i \int_{t_{i-1}}^{t_i} p(t - \theta) d\theta + u_n \int_{t_{n-1}}^t p(t - \theta) d\theta \\
 &= \sum_{i=1}^{n-1} u_i \int_{t-t_i}^{t-t_{i-1}} p(s) ds + u_n \int_0^{t-t_{n-1}} p(s) ds \\
 &\quad (s = t - \theta) \\
 &= \sum_{i=1}^{n-1} u_i [P(t - t_{i-1}) - P(t - t_i)] + u_n P(t - t_{n-1}) \\
 &= \sum_{i=1}^n (\Delta u_i) P(t - t_{i-1}), \tag{13}
 \end{aligned}$$

where $\Delta u_i = u_i - u_{i-1}$ and $u_i = u(i \Delta t)$. Let $t = t_n = n \Delta t$, then

$$y(n \Delta t) = \sum_{i=1}^n (\Delta u_i) P((n - i + 1)\Delta t). \tag{14}$$

Introducing the following variables: $P_i = P(i \Delta t)$, $y_i = y(i \Delta t)$, $\Delta \mathbf{u} = [\Delta u_1 \ \Delta u_2 \ \dots \ \Delta u_n]^T$, $\mathbf{u} = [u_1 \ u_2 \ \dots \ u_n]^T$, $\mathbf{y} = [y_1 \ y_2 \ \dots \ y_n]^T$, the output and input of the dynamic system of Eq. (14) can be formulated as $\mathbf{y} = \mathbf{P} \Delta \mathbf{u}$ and $\mathbf{u} = \mathbf{E} \Delta \mathbf{u}$, where

$$\mathbf{P} = \begin{bmatrix} P_1 & 0 & \dots & \dots & 0 \\ P_2 & P_1 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & P_1 & 0 \\ P_n & \dots & \dots & P_2 & P_1 \end{bmatrix}, \tag{15}$$

and

$$\mathbf{E} = \begin{bmatrix} 1 & 0 & \dots & \dots & 0 \\ 1 & 1 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & 1 & 0 \\ 1 & \dots & \dots & 1 & 1 \end{bmatrix}. \tag{16}$$

If the time interval, Δt , is sufficiently small, the integral in Eq. (12) can be approximated using the trapezoid sum, and the

optimal control problem of Eq. (12) can be converted to the following standard least squares minimization problem:

$$\begin{aligned}
 \min_{\Delta \mathbf{u}} \quad & J = \|\mathbf{A} \Delta \mathbf{u} - \mathbf{b}\|^2 \\
 \text{s.t.} \quad & \mathbf{C} \Delta \mathbf{u} \leq \mathbf{d}, \tag{17}
 \end{aligned}$$

where $\mathbf{A} = \begin{bmatrix} \mathbf{P} \\ \varepsilon \mathbf{E} \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} \mathbf{e} \\ \varepsilon \mathbf{e} \end{bmatrix}$, $\mathbf{C} = \begin{bmatrix} \mathbf{E} \\ -\mathbf{E} \end{bmatrix}$, $\mathbf{d} = \begin{bmatrix} u_{\max} \mathbf{e} \\ -u_{\min} \mathbf{e} \end{bmatrix}$, and $\mathbf{e} = [1 \ 1 \ \dots \ 1]^T$. The constrained optimization problem of Eq. (17) constitutes a very high-order problem (infinite-dimensional when $n = \infty$) owing to the fact that the time interval is defined from zero to infinity. However, one can formulate the following finite-horizon constrained optimal control problem to obtain a computationally efficient formulation:

$$\begin{aligned}
 \min_{u(t)} \quad & J = \int_0^{T_f} (y(t) - y_f)^2 dt + \varepsilon^2 \int_0^{T_f} (u(t) - u_f)^2 dt \\
 \text{s.t.} \quad & y(t) = \int_{\infty}^t u(\theta) p(t - \theta) d\theta, \\
 & u_{\min} \leq u(t) \leq u_{\max} \tag{18}
 \end{aligned}$$

provided that $u_{\text{opt}}(t) = y_{\text{opt}}(t) = u_f = y_f = 1$ as $t \geq T_f$, or $\int_{T_f}^{\infty} (y(t) - y_f)^2 dt \int_{T_f}^{\infty} (u(t) - u_f)^2 dt = 0$. In other words, the fact the process will approach the steady-state for a sufficiently large T_f is taken into account. A representative example will be used to show this behavior in the case study section.

Remark 1. Eq. (14) can be alternatively formulated in the following way:

$$y(n \Delta t) = \sum_{i=1}^n u_i \{P((n - i + 1)\Delta t) - P((n - i)\Delta t)\} \tag{19}$$

or $\mathbf{y} = \Delta \mathbf{P} \mathbf{u}$, where

$$\Delta \mathbf{P} = \begin{bmatrix} P_1 & 0 & \dots & \dots & 0 \\ P_2 - P_1 & P_1 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & P_1 & 0 \\ P_n - P_{n-1} & \dots & \dots & P_2 - P_1 & P_1 \end{bmatrix}. \tag{20}$$

The equivalent least squares minimization problem of Eq. (12) will have then the following form:

$$\begin{aligned}
 \min_{\mathbf{u}} \quad & J = \|\mathbf{A} \mathbf{u} - \mathbf{b}\|^2 \\
 \text{s.t.} \quad & u_{\min} \leq u(t) \leq u_{\max}, \tag{21}
 \end{aligned}$$

where $\mathbf{A} = \begin{bmatrix} \Delta \mathbf{P} \\ \varepsilon \mathbf{I} \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} \mathbf{e} \\ \varepsilon \mathbf{e} \end{bmatrix}$.

Remark 2. Uniformly distributed time intervals are used in the formulation of the analysis of process dynamics and optimal control in this work. When the input is actually comprised of discretized signals with non-uniform time intervals, the output can still be calculated using Eq. (14) or (19). However, numerical integration might be involved in solving Eq. (12).

Remark 3. Generally speaking, the formulation of the optimal control problem using Eq. (21) is superior to the one using Eq. (17) from a viewpoint of computational time. Solving Eq. (17) or (21) involves the operation of high-dimensional matrices and vectors. However, because only linear algebraic operations are involved, the solution of Eq. (21) is very fast (i.e., it takes only several CPU seconds in general to solve Eq. (21) using Matlab on a Pentium 1.7 GHz computer if the dimension of matrix ΔP is no more than 600). Given the fact that the matrix A in Eq. (17) or (21) is not square and the number of rows ($2n$) is larger than the number of columns (n), an accelerated algorithm can be developed based on the concept of singular value decomposition. Let $A = SVD^T$ be the singular value decomposition of matrix A , where S , V and D^T are $2n \times 2n$, $2n \times n$, and $n \times n$ matrices, respectively. Then the norm

$$\begin{aligned} \|Au - b\|^2 &= \|SVD^T u - b\|^2 \\ &= \|VD^T u - S^{-1}b\|^2 \\ &= \left\| \begin{bmatrix} V_r D^T u \\ \mathbf{0} \end{bmatrix} - \begin{bmatrix} b_r \\ b_m \end{bmatrix} \right\|^2 \\ &= \|V_r D^T u - b_r\|^2 + \|b_m\|^2, \end{aligned} \quad (22)$$

where $V = \begin{bmatrix} V_r \\ \mathbf{0} \end{bmatrix}$ in which V_r is a n by n matrix, $S^{-1}b = \begin{bmatrix} b_r \\ b_m \end{bmatrix}$, in which b_r and b_m are both $n \times 1$ vectors. Because b_m is independent of u , a least squares problem equivalent to Eq. (21) can be formulated as follows:

$$\begin{aligned} \min_u \quad & J = \|A_r u - b_r\|^2 \\ \text{s.t.} \quad & u_{\min} \leq u(t) \leq u_{\max}, \end{aligned} \quad (23)$$

where $A_r = V_r D^T$. A representative example will be used to compare the results and solution times using these two least squares representations in the case studies section.

Remark 4. If there are no input constraints present, an explicit analytic solution to Eq. (21) can be obtained by calculating $\partial J / \partial u = 0$, and the solution is

$$\begin{aligned} u_{\text{opt}} &= (A^T A)^{-1} A^T b \\ &= (\Delta P^T \Delta P + \varepsilon^2 I)^{-1} (\Delta P^T + \varepsilon^2 I) e. \end{aligned} \quad (24)$$

Solving such an equation is instantaneous on a modern computer. As will be illustrated in an example in the case studies section, the solution to Eq. (24) is the same as the one obtained from LQR if a state-space model is available.

Remark 5. The above development covered transition processes where the key component to be controlled does not take place in any reactions. When the key component participates in first-order reactions in an isothermal reactor, a similar approach can be applied with the introduction of a modified cumulative residence time distribution. Again, let $p(t)$ be the residence time distribution. The amount of the molecules that enter the reactor at time $t - \theta$ and spend time θ in the reactor before exiting is given by $\eta(\theta)u(t - \theta)p(\theta) d\theta$, where η is the fraction of the molecule that is not consumed by the reaction. Therefore,

the effluent concentration $y(t)$ is the convolution of $u(\theta)$ and $p(\theta)$, or

$$y(t) = \int_0^\infty \eta(\theta)u(t - \theta)p(\theta) d\theta. \quad (25)$$

For a first-order reaction, it can be readily derived that $\eta(\theta) = e^{-k\theta}$, where k is the reaction constant. Therefore, a modified cumulative residence time distribution can be calculated as follows:

$$\begin{aligned} P_m(t) &= \int_0^t e^{-k\theta} p(\theta) d\theta \\ &= e^{-kt} P(t) + k \int_0^t e^{-k\theta} p(\theta) d\theta, \end{aligned} \quad (26)$$

where $P(t)$ is the cumulative residence time distribution defined as $P(t) = \int_0^t p(\theta) d\theta$. However, due to the first-order reaction, it is possible that the output is not equal to the input under steady-state, or $K = \lim_{t \rightarrow \infty} P_m(t) \neq 1$. Then, a similar approach can be followed, provided that $P(t)$ and $\begin{bmatrix} e \\ \varepsilon e \end{bmatrix}$ are replaced by $P_m(t)$ and $\begin{bmatrix} K e \\ \varepsilon e \end{bmatrix}$ in Eq. (21).

Remark 6. If the species to be controlled participates in any reaction that is not of first-order, the residence time distribution alone (macromixing) might not be sufficient to describe the conversion rate due to the micromixing effect (see Fogler, 1998 for examples). In such cases, the concentration evolution is nonlinear and further development of the proposed approach is necessary. The reader may refer to Li and Christofides (accepted for publication) for a state-space approach to the optimal control of diffusion–convection–reaction processes with nonlinear reaction terms and to Dubljevic and Christofides (2006) and Dubljevic et al. (2005, 2006) for results on predictive control of linear/nonlinear parabolic PDEs using reduced-order models.

4. Case studies

As presented in the previous section, the solution of the optimal control trajectory is derived from the cumulative residence time distribution function $P(t)$ instead of the process dynamic equations (state-space model). Therefore, the proposed approach is applicable to both lumped parameter and distributed parameter processes provided the assumptions stated above on the properties of the key component are satisfied. In this section, we use two examples: a series of two CSTRs and a tubular reactor with dispersion and first-order reaction to demonstrate the applicability and evaluate the effectiveness of the proposed method.

4.1. Two CSTRs in series

In this section, we consider two CSTRs in series example because its dynamic behavior can be described in the form of a state-space model and a comparison between different control strategies can be easily made, which include LQR, PI control and the input/output approach proposed in this work. Let u be the concentration of the key component at the entrance of the

reactor, x_1 and x_2 be the concentrations at the exit of the first and the second CSTR, respectively, and τ_1 and τ_2 be the time constants of the first and the second CSTR, respectively. The dynamic evolution of the concentrations can be described by the following state-space model:

$$\begin{aligned} \dot{\tilde{x}} &= A\tilde{x} + b\tilde{u}, \\ \tilde{y} &= c\tilde{x}, \end{aligned} \tag{27}$$

where $\tilde{x} = [x_1 \ x_2]^T$, $A = \begin{bmatrix} -1/\tau_1 & 0 \\ 1/\tau_2 & -1/\tau_2 \end{bmatrix}$, $b = [1/\tau_1 \ 0]^T$, $c = [0 \ 1]$. The tilde sign represents the nominal value of the variable, e.g., $\tilde{x} = x - x_f$, $\tilde{u} = u - u_f$ and $\tilde{y} = y - y_f$. The following initial condition is used $\tilde{x}(0) = [-1 \ -1]^T$ and the following parameter values are used in the simulations: $\tau_1 = \tau_2 = 10$, $u_{\min} = 0$, and $u_{\max} = 5$.

The LQR problem is to minimize the functional

$$\min_{\tilde{u}(t)} J = \int_0^\infty (\tilde{x}^T Q \tilde{x} + \tilde{u}^T R \tilde{u}) dt, \tag{28}$$

and its solution is given by the state feedback law: $\tilde{u} = -k\tilde{x}$, where $k = R^{-1}b^T S$, and S is determined by the Riccati equation:

$$A^T S + SA - SbR^{-1}b^T S + Q = 0. \tag{29}$$

To compare the proposed optimal control method and the standard LQR technique, $Q = c^T c = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$ and $R = \varepsilon^2$ to match the cost function in Eq. (12) and the one in Eq. (28). First the cumulative residence time distribution of the two CSTRs in series is determined as $P(t) = 1 - (1 + t/10) \exp(-t/10)$, which is shown in Fig. 1. The optimization problem is then solved using both control techniques with two different weight settings ($R = \varepsilon^2 = 1$ and $R = \varepsilon^2 = 0.01$, respectively). The closed-loop system under a PI controller with Ziegler–Nichols tuning rules as well as the open-loop system (corresponding to $u(t) \equiv u_f = 1$) is also provided for a comparison.

The dynamic behavior of the closed-loop system under three different control settings as well as the open-loop system is

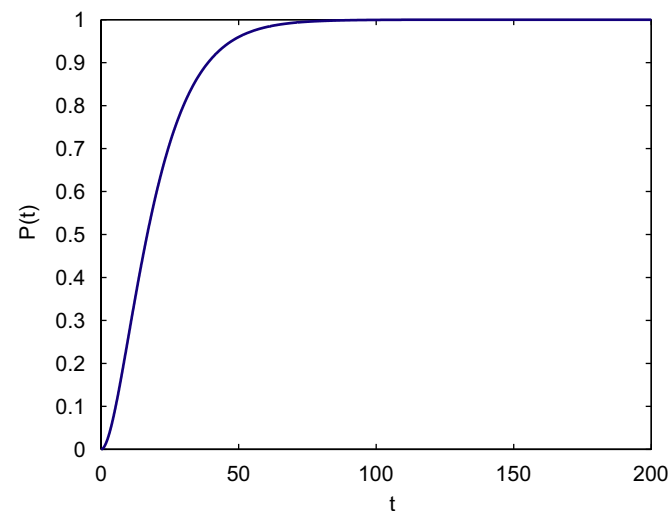


Fig. 1. Cumulative residence time distribution of the two CSTRs in series example.

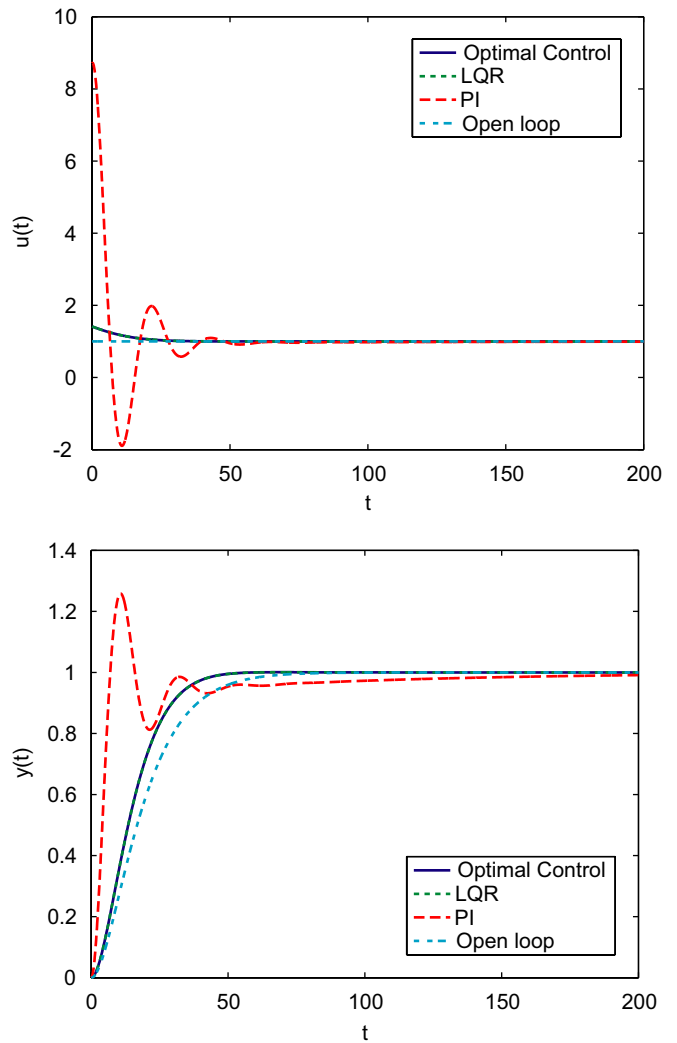


Fig. 2. Profiles of the manipulated input and the controlled output in the two CSTRs in series example using different control schemes. $R = \varepsilon^2 = 1$.

shown in Figs. 2 and 3. In the first case, where $R = \varepsilon^2 = 1$, the closed-loop system demonstrates essentially the same dynamic behavior using the input/output approach proposed in this work as the one using LQR (see Fig. 2). However, in the second case, where $R = \varepsilon^2 = 0.01$, it is found that the manipulated input solved by LQR does not satisfy the input constraints (see Fig. 3). In both cases, the PI controller is not able to handle the input constraints. This example clearly shows that the input/output optimal control scheme proposed in this work is advantageous over both LQR and PI with respect to the satisfaction of input constraints and the minimization of control action. If there are no constraints present, or the constraints, while present, are incidentally satisfied by the manipulated input solved using LQR, it is expected that the proposed optimal control yields the same result as the LQR. This observation is not surprising since both techniques are used to solve the same problem.

As mentioned earlier, a suitable horizon length T_f is required to obtain the optimal solution. Generally speaking, a larger T_f will yield a more accurate solution. However, the computational

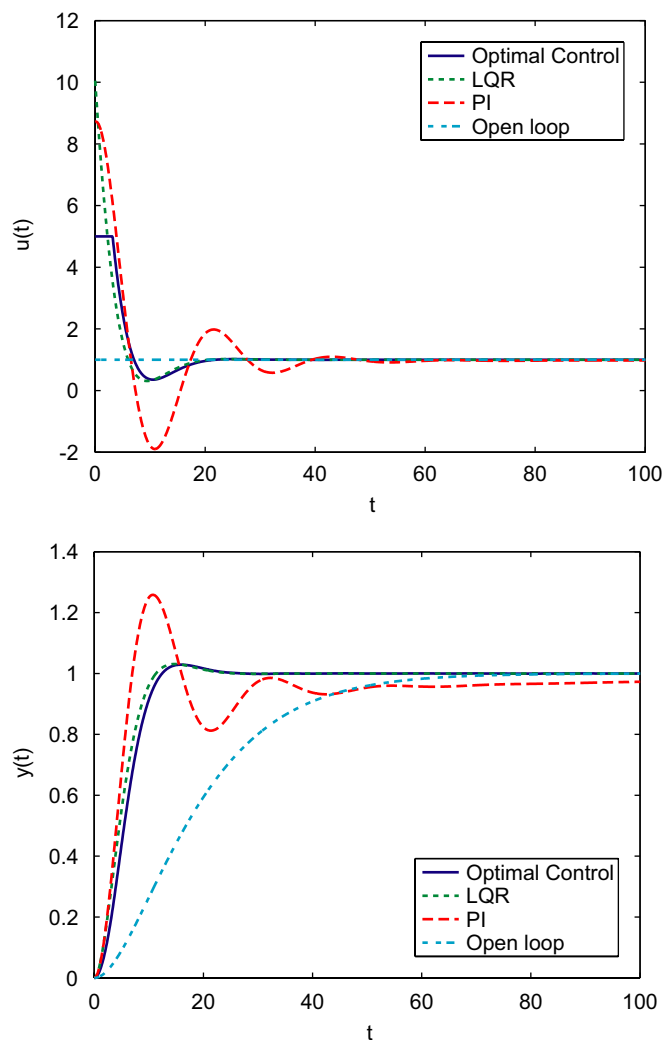


Fig. 3. Profiles of the manipulated input and the controlled output in the two CSTRs in series example using different control schemes. $R = \varepsilon^2 = 0.01$.

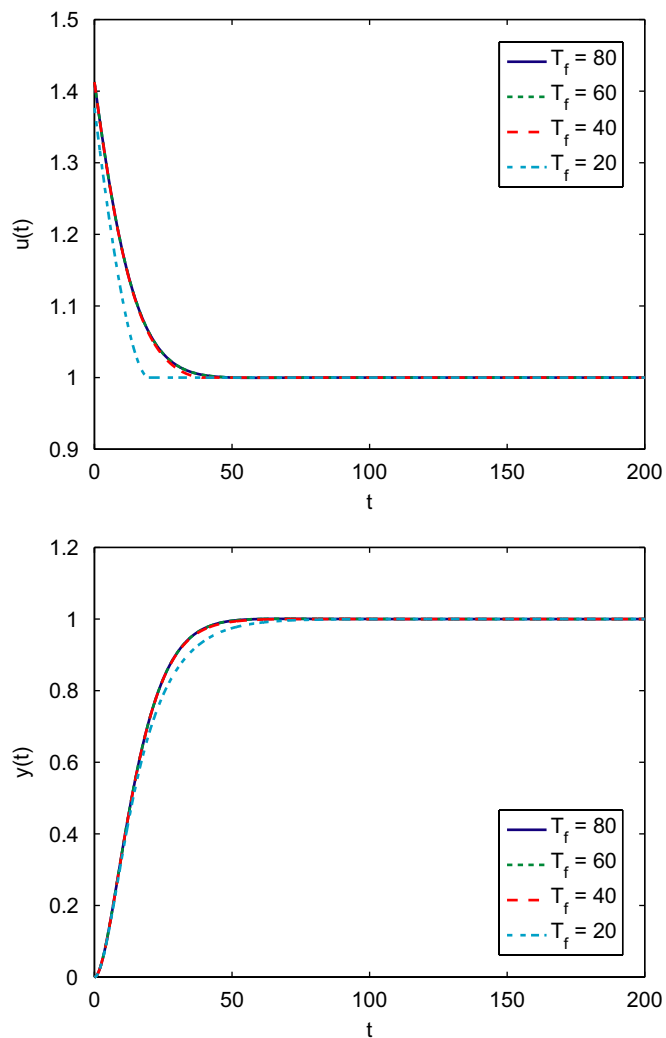


Fig. 4. Profiles of the manipulated input and the controlled output in the two CSTRs in series example under constrained optimal control—effect of different horizon lengths. $R = \varepsilon^2 = 1$.

load will increase accordingly. To demonstrate the influence of the horizon length on the “optimal” solution, the optimal control problem is solved using four different horizon lengths with $R = 1$ and the results are shown in Fig. 4. When $T_f = 20$, the obtained solution is not optimal. As T_f increases, the solution becomes closer to the optimal solution. When $T_f \geq 60$, a further increase in the horizon length has negligible effect on the solution. To guarantee an optimal solution is obtained, the T_f has to be picked such that the process is close to the steady-state at $t = T_f$.

Also, the solution time and accuracy are dependent on the time interval Δt besides the horizon length T_f . Because the dimension of the matrix ΔP is $T_f/\Delta t$, a larger Δt leads to lower computational cost if T_f is fixed. However, the approximation of the integral in Eq. (12) using the linear summation in Eq. (17) or (21) might be less accurate when Δt is not small enough. To study the effect of Δt , the same optimization problem is solved using different time intervals with $R = 0.01$ and

the results are shown in Fig. 5. It is seen clearly that the input profile solved using a large time interval (i.e., $\Delta t = 2$) is different from the one using a short time interval (e.g., $\Delta t = 0.1$). The output profiles are surprisingly close to each other probably because the weight to the output is significantly larger than the input and a near-optimal solution (obtained using large intervals) has very similar output to the one in the actual optimal solution (obtained using small enough intervals). This conclusion may not hold when the time intervals become sufficiently large.

Finally, the same constrained optimal control problem is also solved using a reduced least squares model (Eq. (23)) with $R = 0.01$ (see Remark 3) and a comparison with the original least squares model (Eq. (21)) is shown in Fig. 6. It can be seen that the same profiles of the manipulated input and controlled output are obtained with the reduced model. However, the CPU time decreases from 5.3 to 3.0 s on a 1.7 GHz computer, resulting in a 40% increase in computational efficiency.

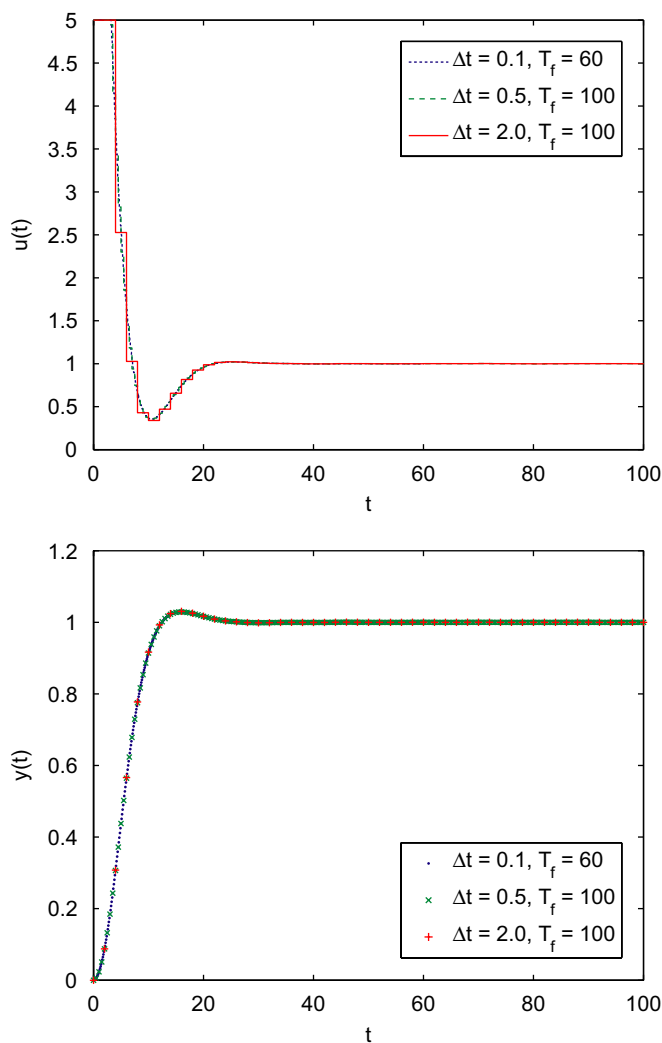


Fig. 5. Profiles of the manipulated input and the controlled output in the two CSTRs in series example under constrained optimal control—effect of different time intervals. $R = \varepsilon^2 = 0.01$.

4.2. A tubular reactor with axial dispersion and first-order reaction

Consider a tubular flow reactor with axial dispersion and a first-order reaction $A \rightarrow B$. The evolution of the concentration of component A is described by a linear parabolic PDE subject to the so-called Danckwerts (1953) boundary conditions:

$$\frac{\partial U(z, t)}{\partial t} = -v \frac{\partial U(z, t)}{\partial z} + D \frac{\partial^2 U(z, t)}{\partial z^2} - kU(z, t)$$

$$\text{s.t. } vU(0^-, t) = uU(0^+, t) - D \left. \frac{\partial U(z, t)}{\partial z} \right|_{z=0^+},$$

$$\left. \frac{\partial U(z, t)}{\partial z} \right|_{z=L} = 0, \quad (30)$$

where $u(t) = U(0^-, t)$ is the inlet concentration, $y(t) = U(L, t)$ is the outlet concentration, t is the time, v is the fluid velocity in the reactor, L is the length of the reactor and D is the dispersion

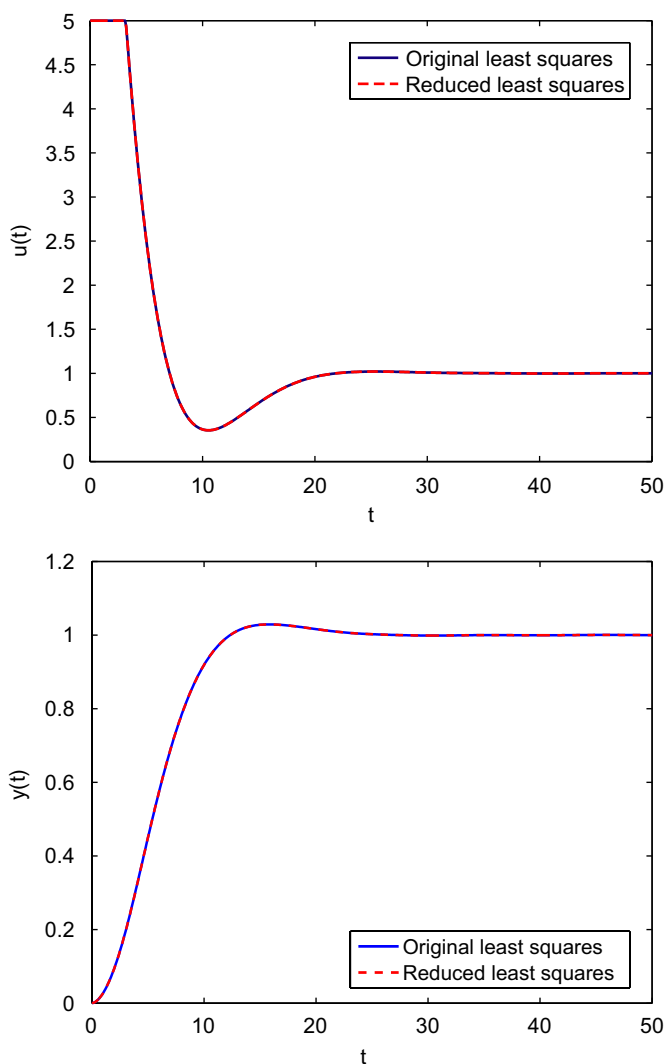


Fig. 6. Profiles of the manipulated input and the controlled output in the two CSTRs in series example under constrained optimal control—comparison between full-order least squares formulation and reduced-order least squares formulation. $R = \varepsilon^2 = 0.01$.

coefficient. The key parameters for this problem are the Peclet number $Pe = vL/D = 10$, the characteristic time $\tau = L/v = 5$ and the Damkohler number $Da = kL/v = 1$.

The dynamics of such a process can be solved by numerical methods based on eigenfunction expansion (e.g., Galerkin's method) and the optimal control trajectory can be calculated based on the reduced-order state-space model derived from the high-order numerical discretization and Galerkin projection (see Li and Christofides, accepted for publication, for details). In the present work, the optimal transition control problem of Eq. (12) is solved using the input/output approach based on the concept of residence time distribution. In particular, the cumulative residence time distribution is first determined by applying an input step change to the high-dimensional state-space model derived from Eq. (30) (note that for more complicated problems, CFD simulations and particle tracing might be used). The calculated $P(t)$ and $P_m(t)$ are shown in Fig. 7. Clearly,

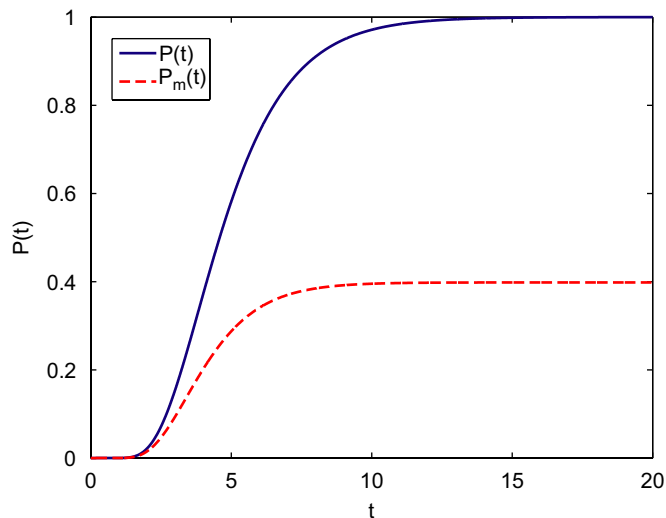


Fig. 7. Profiles of the cumulative residence distribution ($P(t)$) and the modified residence time distribution ($P_m(t)$) for the tubular reactor example.

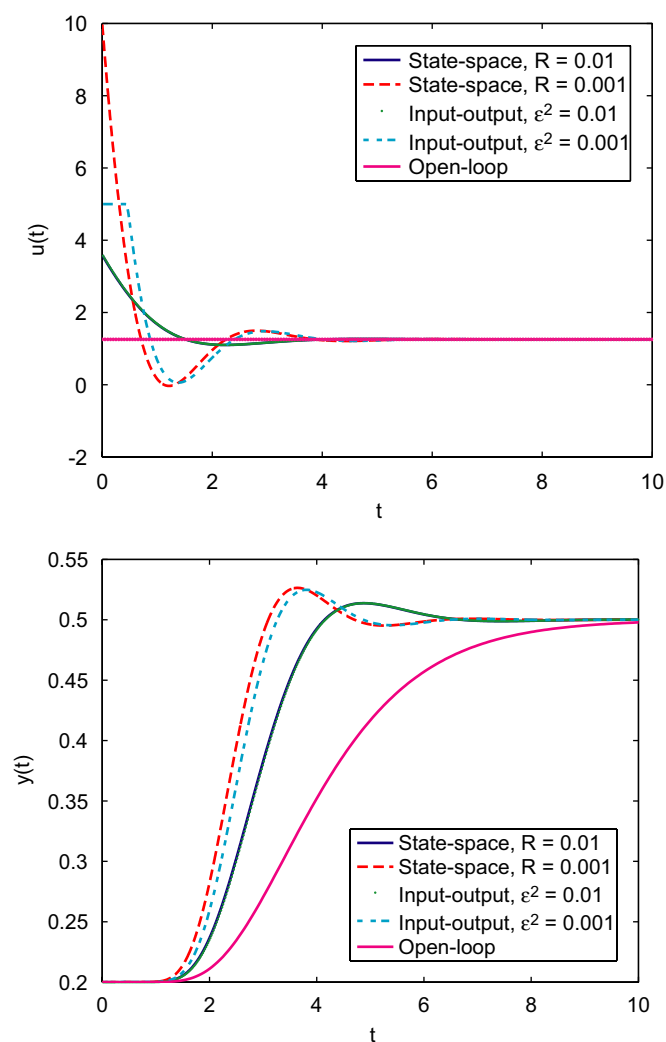


Fig. 8. Comparison of trajectories of the controlled output and the manipulated input solved using the input/output approach and the state-space approach for the tubular reactor example.

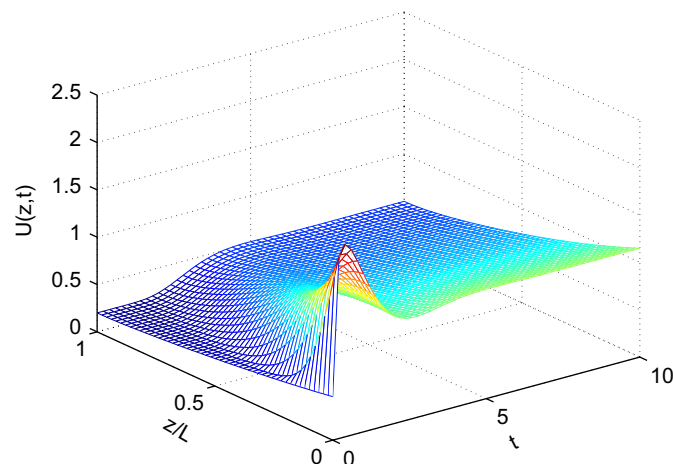


Fig. 9. Profile of the concentration in the closed-loop system with the control action derived using the input/output approach ($\varepsilon^2 = 0.01$) for the tubular reactor example.

$\lim_{t \rightarrow \infty} P_m(t) \neq 1$ due to the presence of the chemical reaction. The optimal control problem is solved with constraints $0 \leq u(t) \leq 5$ using the input/output approach. In the calculation, the control horizon length T_f is chosen to be 10 and the dimension of the matrix $\Delta P_m(T_f/\Delta t)$ is chosen to be 500. A comparison of the input/output approach and the state-space approach (with LQR) is shown in Fig. 8. Clearly, when $R = \varepsilon^2 = 0.01$, in which case the input constraints are coincidentally satisfied using the LQR technique, both approaches yield the same trajectories of the controlled output and the manipulated input. However, when $R = \varepsilon^2 = 0.001$, in which case the LQR technique leads to input constraint violation, the input/output approach computes an optimal input trajectory which satisfies the input constraints. Finally, the control action calculated from the input/output approach with $R = \varepsilon^2 = 0.01$ is implemented on a high-dimensional approximation of the parabolic PDE and the resulting closed-loop concentration profile is shown in Fig. 9.

5. Summary

An optimal control scheme for regulating the concentration transition in a class of distributed chemical reactors was proposed in this work. The control problem formulation was based on the concept of the cumulative residence time distribution of the chemical reactor which might be determined through CFD simulations or experimental measurements. The optimal control action was calculated by solving a finite-dimensional least squares problem with inequality constraints. The advantages of the proposed method over PI control and LQR control were demonstrated through two illustrative examples.

Acknowledgments

Financial support from NSF, CTS-0325246, is gratefully acknowledged. The authors would also like to thank the anonymous reviewers for valuable comments and suggestions.

References

- Armaou, A., Christofides, P.D., 1999. Nonlinear feedback control of parabolic PDE systems with time-dependent spatial domains. *Journal of Mathematical Analysis and Applications* 239, 124–157.
- Armaou, A., Christofides, P.D., 2000. Wave suppression by nonlinear finite-dimensional control. *Chemical Engineering Science* 55, 2627–2640.
- Armaou, A., Christofides, P.D., 2002. Dynamic optimization of dissipative PDE systems using nonlinear order reduction. *Chemical Engineering Science* 57, 5083–5114.
- Baker, J., Christofides, P.D., 2000. Finite dimensional approximation and control of non-linear parabolic PDE systems. *International Journal of Control* 73, 439–456.
- Baker, J., Armaou, A., Christofides, P.D., 2000. Nonlinear control of incompressible fluid flow: application to Burgers' equation and 2D channel flow. *Journal of Mathematical Analysis and Applications* 252, 230–255.
- Balsa-Canto, E., Alonso, A.A., Banga, J.R., 2004. Reduced-order models for nonlinear distributed process systems and their application in dynamic optimization. *Industrial and Engineering Chemistry Research* 43, 3353–3363.
- Balsa-Canto, E., Alonso, A.A., Banga, J.R., 2005. Dynamic optimization of complex distributed process systems. *Chemical Engineering Research Design* 83, 724–729.
- Bendersky, E., Christofides, P.D., 2000. Optimization of transport-reaction processes using nonlinear model reduction. *Chemical Engineering Science* 55, 4349–4366.
- Cervantes, A.M., Tonelli, S., Brandolin, A., Bandoni, J.A., Biegler, L.T., 2002. Large-scale dynamic optimization for grade transitions in a low density polyethylene plant. *Computers and Chemical Engineering* 26, 227–237.
- Christofides, P.D., 2001. *Nonlinear and Robust Control of PDE Systems: Methods and Applications to Transport-Reaction Processes*. Birkhäuser, Boston, USA.
- Christofides, P.D., Daoutidis, P., 1997. Finite-dimensional control of parabolic PDE systems using approximate inertial manifolds. *Journal of Mathematical Analysis and Applications* 216, 398–420.
- Danckwerts, P.V., 1953. Continuous flow systems: distribution of residence times. *Chemical Engineering Science* 2, 1–13.
- Dubljevic, S., Christofides, P.D., 2006. Predictive control of parabolic PDEs with boundary control actuation. *Chemical Engineering Science* 61, 6239–6248.
- Dubljevic, S., Mhaskar, P., El-Farra, N.H., Christofides, P.D., 2005. Predictive control of transport-reaction processes. *Computers and Chemical Engineering* 29, 2335–2345.
- Dubljevic, S., Mhaskar, P., El-Farra, N.H., Christofides, P.D., 2006. Predictive control of parabolic PDEs with state and control constraints. *International Journal of Robust Nonlinear Control* 16, 749–772.
- Ekambara, K., Joshi, J.B., 2004. Axial mixing in laminar pipe flows. *Chemical Engineering Science* 59, 3929–3944.
- Fogler, H.S., 1998. *Elements of Chemical Reaction Engineering*. third ed. Prentice-Hall, New York.
- Graham, M.D., Kevrekidis, I.G., 1996. Alternative approaches to the Karhunen–Loève decomposition for model reduction and data analysis. *Computers and Chemical Engineering* 20, 495–506.
- Graham, W.R., Peraire, J., Tang, K.Y., 1999. Optimal control of vortex shedding using low-order models. Part I open-loop model development. *International Journal for Numerical Methods in Engineering* 44, 945–972.
- Groetsch, M., Gundermann, M., Mangold, M., Kienle, A., Sundmacher, K., 2006. Development and experimental investigation of an extended kalman filter for an industrial molten carbonate fuel cell system. *Journal of Process Control* 16, 985–992.
- Gunjal, P.R., Ranade, V.V., Chaudhari, R.V., 2003. Liquid distribution and RTD in trickle bed reactors: experiments and CFD simulations. *Canadian Journal of Chemical Engineering* 81, 821–830.
- Harris, C.K., Roekaerts, D., Rosendal, F.J.J., Buitendijk, F.G.J., Daskopoulos, Ph., Vreenegoor, A.J.N., Wang, H., 1996. Computational fluid dynamics for chemical reactor engineering. *Chemical Engineering Science* 51, 1569–1594.
- Heibel, A.K., Scheenen, T.W.J., Heiszwolf, J.J., Van As, H., Kapteijn, F., Moulijn, J.A., 2001. Gas and liquid phase distribution and their effect on reactor performance in the monolith film flow reactor. *Chemical Engineering Science* 56, 5935–5944.
- Heibel, A.K., Lebens, P.J.M., Middelhoff, J.W., Kapteijn, F., Moulijn, J., 2005. Liquid residence time distribution in the film flow monolith reactor. *A.I.Ch.E. Journal* 51, 122–133.
- Li, M., Christofides, P.D., 2005. Multi-scale modeling and analysis of HVOF thermal spray process. *Chemical Engineering Science* 60, 3649–3669.
- Li, M., Christofides, P.D., 2006. Computational study of particle in-flight behavior in the HVOF thermal spray process. *Chemical Engineering Science* 61, 6540–6552.
- Li, M., Christofides, P.D., Optimal control of diffusion–convection–reaction processes using reduced-order models. *Computers and Chemical Engineering*, accepted for publication.
- Li, M., Sopko, J.F., McCamy, J.W., 2006. Computational fluid dynamic modeling of tin oxide deposition in an impinging chemical vapor deposition reactor. *Thin Solid Films* 515, 1400–1410.
- Lo, D.P., Ray, W.H., 2006. Dynamic modeling of polyethylene grade transitions in fluidized bed reactors employing nickel–diimine catalysts. *Industrial and Engineering Chemistry Research* 45, 993–1008.
- McAuley, K.B., MacGregor, J.F., 1992. Optimal grade transitions in a gas-phase poly-ethylene reactor. *A.I.Ch.E. Journal* 38, 1564–1576.
- Park, H.M., Lee, M.W., 2000. Control of Navier-Stokes equations by means of mode reduction. *International Journal for Numerical Methods in Fluids* 33, 535–557.
- Raja, L.L., Kee, R.J., Serban, R., Petzold, L.R., 2000. Computational algorithm for dynamic optimization of chemical vapor deposition processes in stagnation flow reactors. *Journal of the Electrochemical Society* 147, 2718–2726.
- Rigopoulos, S., Jones, A., 2003. A hybrid CFD-reaction engineering framework for multiphase reactor modelling: basic concept and application to bubble column reactors. *Chemical Engineering Science* 58, 3077–3089.
- Trier, W., 1987. *Glass Furnaces: Design Construction and Operation*. Society of Glass Technology, Sheffield, UK.
- Varshney, A., Armaou, A., 2006. Optimal operation of GaN thin film epitaxy employing control vector parametrization. *A.I.Ch.E. Journal* 52, 1378–1391.