

Research Article

Data-Driven Approximated Optimal Control for Chemical Processes with State and Input Constraints

Mingfang He 

School of Computer and Information Engineering, Central South University of Forestry and Technology, Changsha 410004, China

Correspondence should be addressed to Mingfang He; t20162306@csuft.edu.cn

Received 13 May 2019; Revised 7 July 2019; Accepted 19 September 2019; Published 28 October 2019

Academic Editor: Pietro De Lellis

Copyright © 2019 Mingfang He. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Input and state constraints widely exist in chemical processes. The optimal control of chemical processes under the coexistence of inequality constraints on input and state is challenging, especially when the process model is only partially known. The objective of this paper is to design an applicable optimal control for chemical processes with known model structure and unknown model parameters. To eliminate the barriers caused by the hybrid constraints and unknown model parameters, the inequality state constraints are first transformed into equality state constraints by using the slack function method. Then, adaptive dynamic programming (ADP) with nonquadratic performance integrand is adopted to handle the augmented system with input constraints. The proposed approach requires only partial knowledge of the system, i.e., the model structure. The value information of the model parameters is not required. The feasibility and performance of the proposed approach are tested using two nonlinear cases including a continuous stirred-tank reactor (CSTR) example.

1. Introduction

Constraints on input and state commonly exist in chemical processes due to finite capability of the actuators [1], safety limits [2], requirements on product quality, and environmental regulations [3]. For example, in practice, it is required to keep the outlet species concentrations of a chemical reactor in reasonable and stable ranges. In addition, the value of manipulated variables is constrained in a certain range defined by the operating instructions. Therefore, the ability to handle constraints is an essential concern in the control design and synthesis of real chemical processes [4, 5].

The study of constrained optimal control has undergone different stages and can be classified into different categories, e.g., problem transformation methods, Lyapunov function-based methods, state-dependent Riccati equation (SDRE) methods, model predictive control (MPC), and machine learning-based methods, to name a few. Since the 1960s, numerous approaches have been proposed to handle the input and/or state constraints in the optimal control of linear/nonlinear systems. Leitmann [6] and Bryson and

Denham [7] proposed penalty-function technique for the optimal control of systems with state-variable inequality constraints; the solution of the penalized optimal control problem converges to the constrained optimal control problem if the penalty multiplier approaches infinity [8]. Sakawa [9] transformed the optimal control of linear systems with input constraints into an infinite-dimensional nonlinear programming problem by integrating the linear differential equation of the system states. Jacobson and Lele [10] used slack variable method to eliminate the scalar inequality constraints on the state variables. Hager [11] applied the Ritz–Trefftz method in the optimal control of system with both state and control constraints. Vlassenbroeck [12] transformed the state-variable inequality constrained optimal control problem into a parameter optimization problem using Chebyshev series expansion. Lim et al. [13] generalized the separation theorem to the constrained linear-quadratic (LQ) and linear-quadratic-Gaussian (LQG) optimal control problem. Manousiouthakis and Chmielewski [14] approximated the optimal control of nonlinear systems subject to pointwise-in-time inequality constraints in the SDRE framework. El-Farra and Christofides [3] developed a

unified framework for constrained optimal control of nonlinear uncertain systems based on a general state-space Lyapunov approach. Balestrino et al. [15] designed a control Lyapunov R-functions (CLRF) solution for constrained linear systems with input and state constraints. Kiefer et al. [16] first transformed the input and state constraints into the constraints on the outputs and their higher-order derivatives, and then, special saturation functions were utilized to incorporate the constraints into overall control design. Stathopoulos et al. [17] studied the linear-quadratic-regulation (LQR) with input and state constraints. By using proximal algorithms and duality, they decomposed the corresponding quadratic program (QP) into two sub-problems including an infinite-dimensional least squares problem and a simple clipping of an infinite sequence to the nonpositive orthant.

MPC is a standard tool to handle input and state constraints within an optimal control setting [15], especially in industry applications [18], e.g., paper and pulp [19], minerals processing [20], chemical engineering [21–23], renewable energy [24], mechatronics engineering [25], and urban water supply [26]. One problem in the MPC implementation is the high computational load for large-scale or fast-sampling systems. Scokaert and Rawlings [27] solved the constrained LQR problem in a finite-dimensional MPC setting with optimality and stability. Wang and Wan [28] developed a structured neural network to solve the QP problem in a constrained MPC. Bemporad [29] moved the MPC computations offline and proposed a technique to obtain the piecewise explicit optimal control law for both MPC and constrained linear-quadratic-regulation (CLQR). By offline approximation, Pin et al. [30] designed a MPC scheme for constrained nonlinear discrete-time systems which allows coping with discontinuous control laws. Mhaskar [31] combined control Lyapunov function (CLF) and MPC to form a Lyapunov-based MPC approach that guarantees stability and constraint satisfaction from an explicitly characterized set of initial conditions. More detailed review of constrained optimal control could be found in [32–34] and the references therein.

Application of the aforementioned methods relies on the full knowledge of system dynamics. With the increased unavailability of quality raw materials, it is imperative that raw materials of low grade with large variations should be employed in the production in order to maximize the use of resources. In this context, the operation of some chemical processes exhibit complexity in terms of variable dynamic characteristics, strong nonlinearities, heavy coupling, unclear mechanism, and mathematically unmodelable parts. Therefore, a common situation is that a chemical process has various working conditions. The structure of the process can be derived by applying conservative laws and its essential physicochemical mechanisms. However, the model parameters are unknown under some working conditions, e.g., insufficient data samples for model identification and undetermined reaction mechanism. This constitutes challenges to the existing control theory and technology.

In recent years, the development of machine learning supported and gave rise to the emergence of data-driven

constrained optimal control methods and the integration of machine learning and MPC [35]. Chakrabarty et al. [36] used a support-vector machine (SVM) to learn feasible region boundaries for an explicit nonlinear model predictive control (ENMPC). Lin and Zheng [37] devised a reinforcement learning agent to obtain the optimal control strategy for nonlinear systems with inequality constraints on input and state via cycle-by-cycle finite-time optimization. Full knowledge of the system model is still required in this approach. Abu-Khalaf and Lewis [38], Modares et al. [39], Luo et al. [40], Yang et al. [41], Zhang et al. [42], and Zhu et al. [43] used adaptive dynamic programming (ADP) [44–46] to approximate the optimal state-feedback controller for input-constrained nonlinear systems. Fan and Yang [47] applied ADP to approximate the optimal control law for state-constrained nonlinear systems. These ADP-based approaches are model-free or do not need partial knowledge of the system model; however, they provide the optimal control solution for either input-constrained or state-constrained systems. Chi et al. [48] studied the constrained data-driven iterative learning control (ILC) for point-to-point optimal control problem of discrete-time nonlinear systems with input and output constraints.

The aim of this study is to design an approximated optimal control algorithm for continuous-time nonlinear chemical processes with both state and input constraints. The proposed algorithm requires only the structural knowledge of the system model, the value of the model parameters is not needed, which is a common situation in industrial applications. The input and state constraints were handled in a sequential manner. The state constraints were first eliminated by introducing slack functions to form an augmented system without inequality state constraints. Then, a nonquadratic performance integrand was adopted in the ADP framework to account for the input constraints in the augmented system.

The rest of this paper is organized as follows. In Section 2, the constrained optimal control is formulated, some preliminaries are introduced. The constraint handling approach and the approximated optimal control design are described in Section 3. Simulation results are presented and discussed in Section 4, followed by concluding remarks in Section 5.

2. Problem Formulation and Preliminaries

2.1. Formulation of the Nonlinear Constrained Optimal Control Problem. Consider a chemical process described by following continuous-time constrained system:

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) + \mathbf{g}(\mathbf{x})\mathbf{u}, \quad (1)$$

where $\mathbf{x} \in \mathbb{X}$ is the system state and $\mathbf{u} \in \mathbb{U}$ is the manipulated input; $\mathbf{f}(\mathbf{x}) \in \mathbb{R}^n$ and $\mathbf{g}(\mathbf{x}) \in \mathbb{R}^{n \times m}$ are differentiable and Lipschitz continuous functions. \mathbb{X} and \mathbb{U} are compact and contain the equilibrium point, i.e., the origin in their interiors. More specifically, $\mathbb{X} = \{\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n : \mathbf{C}_{\min} \leq \mathbf{C}(\mathbf{x}) \leq \mathbf{C}_{\max}\}$ and $\mathbb{U} = \{\mathbf{u} = (u_1, \dots, u_m) \in \mathbb{R}^m : \mathbf{u}_{\min} \leq \mathbf{u} \leq \mathbf{u}_{\max}\}$ denote the hard constraints on the states and inputs, respectively. $\mathbf{C}(\mathbf{x})$ is a Lipschitz continuous

function of \mathbf{x} , which demonstrates the physical or technical constraints on the system states. \mathbf{C}_{\max} , \mathbf{C}_{\min} , \mathbf{u}_{\max} , and \mathbf{u}_{\min} are the maximum and minimum value vectors of $\mathbf{C}(\mathbf{x})$ and input variables.

For chemical processes described by (1) with system state and input constrained in \mathbb{X} and \mathbb{U} , the aim of optimal control is to find a control policy which minimizes the following performance index:

$$V(\mathbf{x}_0) = \int_0^{t_f} [W(\mathbf{x}(\tau)) + M(\mathbf{u}(\tau))] d\tau, \quad (2)$$

where $V(\mathbf{x}_0)$ is the value function and $W(\cdot)$ and $M(\cdot)$ are positive definite weighting functions of the immediate states and inputs, respectively. $V(\mathbf{x}_0)$ is an overall index of the control performance. A lower value of $V(\mathbf{x}_0)$ indicates better control performance. $W(\mathbf{x})$ evaluates the deviation of the system state with the origin, e.g., the difference between the actual outlet concentration of a chemical reactor with its setting value. $M(\mathbf{u})$ accounts for the control effort to be placed on the process, e.g., the dosage of additive for the reactor to change the outlet concentration. $[0, t_f]$ is the time interval of interest.

2.2. Preliminaries on Optimal Control of Nonlinear Systems. To start with, the preliminaries on optimal control of nonlinear systems are introduced without considering the system and input constraints. When the system is unconstrained, $W(\mathbf{x})$ and $M(\mathbf{u})$ are usually selected as $W(\mathbf{x}) = \mathbf{x}^T \mathbf{Q} \mathbf{x}$ and $M(\mathbf{u}) = \mathbf{u}^T \mathbf{R} \mathbf{u}$, with $\mathbf{Q} \geq 0$ and $\mathbf{R} > 0$, the symmetric weighting matrices. The solution of the unconstrained optimal control problem can be obtained by solving a Hamilton–Jacobi–Bellman (HJB) equation:

$$V_x^T \mathbf{f}(\mathbf{x}) + \mathbf{x}^T \mathbf{Q} \mathbf{x} - \frac{1}{4} V_x^T \mathbf{g}(\mathbf{x}) \mathbf{R}^{-1} \mathbf{g}(\mathbf{x})^T V_x = 0, \quad (3)$$

with the boundary condition $V(0) = 0$. V_x is the partial derivative of the cost function V regarding system state \mathbf{x} . If the HJB equation has an optimal solution V^* , then the optimal control is

$$\mathbf{u}^* = -\frac{1}{2} \mathbf{g}(\mathbf{x})^T V_x^*. \quad (4)$$

The HJB equation involves a nonlinear differential equation whose analytic solution is difficult to obtain. However, for the HJB equation, if a control policy \mathbf{u} has been improved using $V(\mathbf{x})$ to generate a better policy \mathbf{u}' , then $V(\mathbf{x})'$ corresponding to the improved policy \mathbf{u}' can be used to yield an even better policy \mathbf{u}'' . This approach of solving the optimal control policy is called policy iteration, which can gradually converge to an optimal policy and optimal value function.

Therefore, successive approximation methods are developed to iteratively improve the control policy [49]. For system (1), starting from an initial admissible control $\mathbf{u}_0(\mathbf{x})$, two sequences of $\{V_i(\mathbf{x})\}_{i=0}^{\infty}$ and $\{\mathbf{u}_{i+1}(\mathbf{x})\}_{i=0}^{\infty}$ could be generated via the policy iteration (PI) approach defined in [49]:

- (1) For $i \geq 0$, solve the Lyapunov equation (LE) which is linear in V :

$$V_{i_x}^T [\mathbf{f}(\mathbf{x}) + \mathbf{g}(\mathbf{x}) \mathbf{u}_i] + \mathbf{x}^T \mathbf{Q} \mathbf{x} + \mathbf{u}_i^T \mathbf{R} \mathbf{u}_i = 0. \quad (5)$$

- (2) Use the solution V_{i_x} to update the control policy:

$$\mathbf{u}_i^* = -\frac{1}{2} \mathbf{g}(\mathbf{x})^T V_{i_x}^*, \quad (6)$$

and

- (1) $0 \leq V_{i+1}(\mathbf{x}) \leq V_i(\mathbf{x})$
- (2) $\mathbf{u}_i(\mathbf{x})$ is admissible
- (3) if V^* and \mathbf{u}^* exist, $V_i(\mathbf{x}) \rightarrow V^*$ and $\mathbf{u}_i(\mathbf{x}) \rightarrow \mathbf{u}^*$ [50]

The above is the general iterative framework for solving the optimal control problem of nonlinear processes. Different from the unconstrained idea case, for practical chemical processes, the evolution trajectory of the system must obey the constraints. So, the feasible region of state and input is shrunk. These constraints must be taken into account in the optimal control design in order to avoid violating physical or technical limits.

3. Approximated Constrained Optimal Controller Design with Partially Known System Dynamics

In this section, an optimal controller design approach without knowing the full system dynamics is proposed for constrained nonlinear systems. The state constraints are eliminated first by introducing slack functions. Then, the input constraints are managed by using ADP in the optimal control of the resulted augmented system. The iterative solution procedure introduced in Section 2 relies on the precise knowledge of the system dynamics $\mathbf{f}(\mathbf{x})$ and $\mathbf{g}(\mathbf{x})$. In order to eliminate this dependence, in this section, an approximated optimal control is studied and designed.

3.1. Handling State Constraints Using Slack Functions. Consider the state constraint $\mathbf{C}_{\min} \leq \mathbf{C}(\mathbf{x}) \leq \mathbf{C}_{\max}$; it contains the following inequality constraints:

$$\begin{aligned} \mathbf{C}_{\min} - \mathbf{C} &\leq 0, \\ \mathbf{C} - \mathbf{C}_{\max} &\leq 0. \end{aligned} \quad (7)$$

There exists slack function $l_i(\alpha_i(t))$ ([10, 47]), which satisfy, for each state constraint,

$$\eta_i(\mathbf{x}) + l_i(\alpha_i(t)) = 0, \quad (8)$$

where $\eta_i(\mathbf{x})$ could be $(C_i(\mathbf{x}) - C_i^{\max})$, $(C_i^{\min} - C_i(\mathbf{x}))$, or $(C_i(\mathbf{x}) - C_i^{\max})(C_i^{\min} - C_i(\mathbf{x}))$, and according to the structure of $\eta_i(\mathbf{x})$ and the value of C_i^{\min} and C_i^{\max} , the slack function $l_i(\alpha_i) \geq 0$ could be selected as $\gamma_i e^{\alpha_i(t)}$ or $1/2\alpha_i^2$, $\gamma_i > 0$.

By using (8), the inequality state constraints could be transformed into equality state constraints. For example, consider a single inequality constraint:

$$C_i^{\min} - C_i(\mathbf{x}) \leq 0. \quad (9)$$

If

$$l_i(\alpha_i) = \gamma_i e^{\alpha_i(t)}, \quad (10)$$

then γ_i is chosen that

$$\gamma_i = -\eta_i(0). \quad (11)$$

In addition, the relationship between the time derivative of $\alpha_i(t)$ and the input could be obtained by differentiating (8) p_i ($p_i \geq 1$) times until the appearance of input \mathbf{u} :

$$\frac{\partial l_i}{\partial \alpha_i} \dot{\alpha}_i + \eta_{i1}(\mathbf{x}) = 0,$$

$$\left(\frac{\partial^2 l_i}{\partial \alpha_i^2} \dot{\alpha}_i^2 + \frac{\partial l_i}{\partial \alpha_i} \ddot{\alpha}_i \right) + \eta_{i2}(\mathbf{x}) = 0,$$

⋮

$$\frac{\partial l_i}{\partial \alpha_i} \alpha_i^{(p_i)} + \left[l_{i p_i}(\alpha_i) - \frac{\partial l_i}{\partial \alpha_i} \alpha_i^{(p_i)} \right] + \eta_{i p_i}(\mathbf{x}, \mathbf{u}) = 0, \quad (12)$$

where η_{ij} is the j th derivative of $\eta_i(\mathbf{x})$ ($1 \leq j \leq p_i$) and $l_{i p_i}(\alpha_i)$ is the p_i th derivative of $l_i(\alpha_i)$.

Denote the j th order derivative of α_i as α_{ij} , and decompose $\eta_{i p_i}(\mathbf{x}, \mathbf{u})$ as $\eta_{i p_i}(\mathbf{x}, \mathbf{u}) = \Gamma_{i1}(\mathbf{x}) + \Gamma_{i2}(\mathbf{x})\mathbf{u}$. If there exist k ($k \geq 1$) state constraints and N_s ($N_s = \sum_{i=1}^k p_i$) slack variables, then combining (12) with system (1) could form the following augmented system:

$$\dot{\mathbf{x}}_s = \mathbf{f}_s(\mathbf{x}_s) + \mathbf{g}_s(\mathbf{x}_s)\mathbf{u}, \quad (13)$$

where $\mathbf{x}_s = [\mathbf{x}, \alpha_1, \alpha_{11}, \alpha_{12}, \dots, \alpha_{1 p_1}, \dots, \alpha_k, \alpha_{k1}, \alpha_{k2}, \dots, \alpha_{k p_k}]^T$,

$$\mathbf{f}_s(\mathbf{x}_s) = \begin{bmatrix} \mathbf{f}(\mathbf{x}) \\ \alpha_{11} \\ \alpha_{12} \\ \vdots \\ -\left(\frac{\partial l_1}{\partial \alpha_1}\right)^{-1} \left\{ \Gamma_{11}(\mathbf{x}) + \left[l_{1 p_1}(\alpha_1) - \frac{\partial l_1}{\partial \alpha_1} \alpha_1^{(p_1)} \right] \right\} \\ \vdots \\ \alpha_{k1} \\ \alpha_{k2} \\ \vdots \\ -\left(\frac{\partial l_k}{\partial \alpha_k}\right)^{-1} \left\{ \Gamma_{k1}(\mathbf{x}) + \left[l_{k p_k}(\alpha_k) - \frac{\partial l_k}{\partial \alpha_k} \alpha_k^{(p_k)} \right] \right\} \end{bmatrix}, \quad (14)$$

$$\mathbf{g}_s(\mathbf{x}_s) = \begin{bmatrix} \mathbf{g}(\mathbf{x}) \\ 0 \\ 0 \\ \vdots \\ -\left(\frac{\partial l_1}{\partial \alpha_1}\right)^{-1} \Gamma_{12}(\mathbf{x}) \\ \vdots \\ 0 \\ 0 \\ \vdots \\ -\left(\frac{\partial l_k}{\partial \alpha_k}\right)^{-1} \Gamma_{k2}(\mathbf{x}) \end{bmatrix}. \quad (15)$$

From (8), for each i , if $\alpha_i(t)$ exists and is bounded, then the state constraints are not violated. Therefore, regarding the following performance index, if there exists an optimal controller \mathbf{u}^* for system (13) without state constraints, then \mathbf{u}^* is also optimal for the state-constrained system (1):

$$V(\mathbf{x}_0) = \int_0^{t_f} [W_s(\mathbf{x}_s(\tau)) + M(\mathbf{u}(\tau))] d\tau, \quad (16)$$

where $W_s(\mathbf{x}_s) = \mathbf{x}_s^T \mathbf{Q}_s \mathbf{x}_s$ and $\mathbf{Q}_s = \text{diag}(\mathbf{Q}, 0)$ with a dimension $n + N_s$.

3.2. Handling Input Constraints by Using Nonquadratic Performance Integrand. In order to incorporate the control constraints in the optimal control design, the nonquadratic sufficiently differentiable performance integrand could be applied [51]:

$$M(\mathbf{u}) = 2 \int_0^{\mathbf{u}} \mathbf{H}^{-T}(\mathbf{v}) \mathbf{R} d\mathbf{v}, \quad (17)$$

where $\mathbf{H}(\cdot) = [h_1(\cdot), h_2(\cdot), \dots, h_m(\cdot)]^T$ are vectors of bounded, monotonic C^ϵ functions ($\zeta \geq 1$ and $\epsilon \geq 1$) to approximate the hard constraints

$$\mathbf{u}_{\min} \leq \mathbf{H}(\cdot) \leq \mathbf{u}_{\max}. \quad (18)$$

The corresponding Hamiltonian and HJB equations using the nonquadratic performance integrand are

$$V(\mathbf{x}_s) = \left[\mathbf{x}_s^T \mathbf{Q}_s \mathbf{x}_s + \int_0^{\mathbf{u}} \mathbf{H}^{-T}(\mathbf{v}) \mathbf{R} d\mathbf{v} \right] + \lambda^T (\mathbf{f}_s(\mathbf{x}_s) + \mathbf{g}_s(\mathbf{x}_s)\mathbf{u}), \quad (19)$$

$$V_{x_s}^T \left[f(\mathbf{x}_s) - g(\mathbf{x}_s) \mathbf{H} \left(\frac{1}{2} \mathbf{R}^{-1} g^T(\mathbf{x}_s) V_{x_s} \right) \right] + \mathbf{x}_s^T \mathbf{Q}_s \mathbf{x}_s \quad (20)$$

$$+ 2 \int_0^{-\mathbf{H} \left(\frac{1}{2} \mathbf{R}^{-1} g^T(\mathbf{x}_s) V_{x_s} \right)} (\mathbf{H}^{-1}(\mathbf{v}))^T \mathbf{R} d\mathbf{v} = 0,$$

where $\lambda(t) \in \mathbb{R}^n$ is the Lagrange multiplier and $V(0) = 0$.

If the solution $V^*(\mathbf{x})$ of the HJB equation exists, the optimal control law is

$$\mathbf{u}^*(\mathbf{x}) = -\mathbf{H}\left(\frac{1}{2}\mathbf{R}^{-1}\mathbf{g}^T(\mathbf{x}_s)V_{x_s}^*\right), \quad (21)$$

which can be approximated by using a computational approach based on PI [38]:

(1) Choose an initial stabilizing admissible controller $\mathbf{u}_0(\mathbf{x}_s)$.

(2) For $i \geq 0$, solve the following LE for $V(i)$:

$$\begin{aligned} V_{x_s}^T [f_s(\mathbf{x}_s) + g_s(\mathbf{x}_s)\mathbf{u}_i(\mathbf{x}_s)] + \mathbf{x}_s^T \mathbf{Q}_s \mathbf{x}_s \\ + 2 \int_0^{\mathbf{u}_i(\mathbf{x}_s)} (\mathbf{H}^{-1}(\mathbf{v}))^T \mathbf{R} d\mathbf{v} = 0. \end{aligned} \quad (22)$$

(3) Update the control law:

$$\mathbf{u}_{i+1}(\mathbf{x}_s) = -\mathbf{H}\left(\frac{1}{2}\mathbf{R}^{-1}\mathbf{g}^T(\mathbf{x}_s)V_{x_s}\right). \quad (23)$$

3.3. Approximated Optimal Control Design. In Section 3.1 and 3.2, the problems caused by the constraints on the states and inputs are eliminated, respectively. However, the PI denoted by (22) and (23) relies on the full knowledge of the system dynamics. To eliminate this requirement, a computational approach [50] is applied and tailored. For denotation simplicity, in this subsection, \mathbf{x}_s , $f_s(\mathbf{x}_s)$, $g_s(\mathbf{x}_s)$, and $W_s(\mathbf{x}_s)$ are denoted as \mathbf{x} , $f(\mathbf{x})$, $g(\mathbf{x})$, and $W(\mathbf{x})$, respectively.

To start with, decompose \mathbf{u} as

$$\mathbf{u} = \mathbf{u}^{(i)} + \mathbf{v}^{(i)}, \quad (24)$$

then

$$\dot{x}_s = f_s(\mathbf{x}_s) + g_s(\mathbf{x}_s)\mathbf{u}^{(i)} + g_s(\mathbf{x}_s)\mathbf{v}^{(i)}, \quad (25)$$

where \mathbf{u} is the control forced actually on the system and $\mathbf{u}^{(i)}$ is the control policy to be improved iteratively.

Consider a cost function $V^{(i)}$ satisfying (22):

$$\begin{aligned} \dot{V}^{(i)} &= V_{x_s}^{(i)T} \{ [f_s(\mathbf{x}_s) + g_s(\mathbf{x}_s)\mathbf{u}^{(i)}] + g_s(\mathbf{x}_s)\mathbf{v}^{(i)} \} \\ &= -W_s(\mathbf{x}_s) - M(\mathbf{u}^{(i)}) - 2\mathbf{H}^{-T}(\mathbf{u}^{(i+1)})\mathbf{R}\mathbf{v}^{(i)}. \end{aligned} \quad (26)$$

Integrate $\dot{V}^{(i)}$ on a time interval $[t, t + \Delta t]$:

$$\begin{aligned} V^{(i)}(t + \Delta t) - V^{(i)}(t) \\ = \int_t^{t+\Delta t} [-W_s(\mathbf{x}_s) - M(\mathbf{u}^{(i)}) - 2\mathbf{H}^{-T}(\mathbf{u}^{(i+1)})\mathbf{R}\mathbf{v}^{(i)}] d\tau. \end{aligned} \quad (27)$$

Assume $\{\phi_j(\mathbf{x}_s)\}_{j=1}^\infty$ and $\{\psi_j(\mathbf{x}_s)\}_{j=1}^\infty$ are two infinite sequences of linearly independent smooth basis functions on \mathbb{X} , where $\phi_j(0) = 0$ and $\psi_j(0) = 0$ for all $j = 1, 2, \dots$. Approximate the cost function and the control policy as

$$\widehat{V}^{(i)} = \sum_{j=1}^{N_1} \widehat{c}_j^{(i)} \phi_j(\mathbf{x}_s) = \Phi(\mathbf{x}_s)^T \widehat{\mathbf{c}}^{(i)}, \quad (28)$$

where $\widehat{\mathbf{c}}^{(i)} = [\widehat{c}_1^{(i)} \ \widehat{c}_2^{(i)} \ \dots \ \widehat{c}_{N_1}^{(i)}]^T$, $\Phi(\mathbf{x}) = [\phi_1(\mathbf{x}) \ \phi_2(\mathbf{x}) \ \dots \ \phi_{N_1}(\mathbf{x})]^T$, and

$$\widehat{\mathbf{u}}^{(i)} = -\mathbf{H}(\mathbf{R}^{-T}\widehat{\mathbf{Y}}_i(\mathbf{x}_s)), \quad (29)$$

where

$$\widehat{\mathbf{Y}}_i(\mathbf{x}_s) = \widehat{\mathbf{k}}^{(i)}\Psi(\mathbf{x}_s), \quad (30)$$

with

$$\widehat{\mathbf{k}}^{(i)} = \begin{bmatrix} \widehat{k}_{11}^{(i)} & \widehat{k}_{12}^{(i)} & \dots & \widehat{k}_{1N_2}^{(i)} \\ \widehat{k}_{21}^{(i)} & \widehat{k}_{22}^{(i)} & \dots & \widehat{k}_{2N_2}^{(i)} \\ \vdots & \vdots & \ddots & \vdots \\ \widehat{k}_{m1}^{(i)} & \widehat{k}_{m2}^{(i)} & \dots & \widehat{k}_{mN_2}^{(i)} \end{bmatrix}, \quad (31)$$

$$\Psi(\mathbf{x}_s) = [\psi_1(\mathbf{x}_s) \ \psi_2(\mathbf{x}_s) \ \dots \ \psi_{N_2}(\mathbf{x}_s)]^T. \quad (32)$$

In (28)–(32), $N_1 > 0$ and $N_2 > 0$ are two sufficiently large integers. Substitute equations (28)–(32) into (27), and denote t and $t + \Delta t$ as t_k and t_{k+1} :

$$\begin{aligned} &\sum_{j=1}^{N_1} \widehat{c}_j^{(i)} [\phi_j(\mathbf{x}_s(t_{k+1})) - \phi_j(\mathbf{x}_s(t_k))] \\ &- \int_{t_k}^{t_{k+1}} \widehat{\mathbf{k}}^{(i+1)}\Psi(\mathbf{x}_s)(\mathbf{u}_j(\mathbf{x}_s) - \mathbf{u}_j^{(i)}(\mathbf{x}_s)) dt \\ &= - \int_{t_k}^{t_{k+1}} [W_s(\mathbf{x}_s) + M(\mathbf{u}^{(i)})] dt + \varepsilon_{i,k}. \end{aligned} \quad (33)$$

In (33), there are $N_1 + N_2$ variables to be determined in $\widehat{\mathbf{c}}^{(i)}$ and $\widehat{\mathbf{k}}_j^{(i+1)}$ ($j = 1, 2, \dots, m$), considering a sufficient long time sequence $\{t_k\}_{k=0}^L$ with $L \geq N_1 + N_2$; then, according to (33),

$$[\mathbf{I}_\Phi - (\mathbf{I}_{\Psi\mathbf{u}} + \mathbf{I}_{\Psi\mathbf{u}^{(i)}})] \begin{bmatrix} \widehat{\mathbf{c}}^{(i)} \\ \text{vec}(\widehat{\mathbf{k}}^{(i+1)T}) \end{bmatrix} = -\mathbf{I}_W - \mathbf{I}_{M\mathbf{u}^{(i)}} + \Delta_i, \quad (34)$$

where

$$\mathbf{I}_\Phi = \begin{bmatrix} \Phi(\mathbf{x}_s(1))^T - \Phi(\mathbf{x}_s(0))^T \\ \Phi(\mathbf{x}_s(2))^T - \Phi(\mathbf{x}_s(1))^T \\ \vdots \\ \Phi(\mathbf{x}_s(L))^T - \Phi(\mathbf{x}_s(L-1))^T \end{bmatrix}_{L \times N_1}, \quad (35)$$

$$\mathbf{I}_{\Psi\mathbf{u}} = \begin{bmatrix} \int_{t_0}^{t_1} \mathbf{u}^T [I_m \otimes \Psi(\mathbf{x}_s)^T] dt \\ \int_{t_1}^{t_2} \mathbf{u}^T [I_m \otimes \Psi(\mathbf{x}_s)^T] dt \\ \vdots \\ \int_{t_{L-1}}^{t_L} \mathbf{u}^T [I_m \otimes \Psi(\mathbf{x}_s)^T] dt \end{bmatrix}_{L \times mN_2}, \quad (36)$$

$$\mathbf{I}_{\Psi\mathbf{u}^{(i)}} = \begin{bmatrix} \int_{t_0}^{t_1} \mathbf{H} \left(\frac{1}{2} \mathbf{R}^{-1} \hat{k}^{(i)} \Psi(\mathbf{x}_s) \right)^T [I_m \otimes \Psi(\mathbf{x}_s)^T] dt \\ \int_{t_1}^{t_2} \mathbf{H} \left(\frac{1}{2} \mathbf{R}^{-1} \hat{k}^{(i)} \Psi(\mathbf{x}_s) \right)^T [I_m \otimes \Psi(\mathbf{x}_s)^T] dt \\ \vdots \\ \int_{t_{L-1}}^{t_L} \mathbf{H} \left(\frac{1}{2} \mathbf{R}^{-1} \hat{k}^{(i)} \Psi(\mathbf{x}_s) \right)^T [I_m \otimes \Psi(\mathbf{x}_s)^T] dt \end{bmatrix}_{L \times mN_2}, \quad (37)$$

$$\mathbf{I}_W = \begin{bmatrix} \int_{t_0}^{t_1} W(\mathbf{x}_s) dt \\ \int_{t_1}^{t_2} W(\mathbf{x}_s) dt \\ \vdots \\ \int_{t_{L-1}}^{t_L} W(\mathbf{x}_s) dt \end{bmatrix}_{L \times 1}, \quad (38)$$

$$\mathbf{I}_{M\mathbf{u}^{(i)}} = \begin{bmatrix} \int_{t_0}^{t_1} M(\mathbf{u}^{(i)}) dt \\ \int_{t_1}^{t_2} M(\mathbf{u}^{(i)}) dt \\ \vdots \\ \int_{t_{L-1}}^{t_L} M(\mathbf{u}^{(i)}) dt \end{bmatrix}_{L \times 1}, \quad (39)$$

and $\Delta_i = [\varepsilon_{i,1} \ \varepsilon_{i,1} \ \cdots \ \varepsilon_{i,L}]^T \in \mathbb{R}^L$ is the approximation error.

From (34), if

$$\text{rank}(\Theta(\hat{k}^{(i)})) = N_1 + mN_2, \quad (40)$$

where $\Theta(\hat{k}^{(i)}) = [\mathbf{I}_\Phi \ -(\mathbf{I}_{\Psi\mathbf{u}} + \mathbf{I}_{\Psi\mathbf{u}^{(i)}})]$, then $\hat{c}^{(i)}$ and $\hat{k}^{(i+1)}$ can be obtained as

$$\begin{bmatrix} \hat{c}^{(i)} \\ \text{vec}(\hat{k}^{(i+1)^T}) \end{bmatrix} = \left[\Theta(\hat{k}^{(i)})^T \Theta(\hat{k}^{(i)}) \right]^{-1} \Theta(\hat{k}^{(i)})^T \Pi(\hat{k}^{(i)}), \quad (41)$$

where $\Pi(\hat{k}^{(i)}) = [-\mathbf{I}_W - \mathbf{I}_{M\mathbf{u}^{(i)}}]$.

Equation (41) along with (28)–(32) serves as a computational approach to approximate $V^{(i)}$ and generate an improved controller $\mathbf{u}^{(i+1)}$ when an admissible control policy $\mathbf{u}^{(i)}$ is given. By using this approach iteratively, starting from an admissible bounded control $\mathbf{u}^{(0)}$, a control sequence $\{\mathbf{u}^{(i)}(\mathbf{x})\}_{i=0}^{\infty}$ could be obtained, and it satisfies the following theorem.

Theorem 1. Consider system (1) with both state and input constraints, and augment it as (13) using an appropriate number of slack functions; if there exists an optimal controller and Assumption 1-2 hold, then

$$\mathbf{u}^{(i+1)} = -\mathbf{H} \left(\mathbf{R}^{-T} \hat{k}^{(i)} \Psi(\mathbf{x}_s) \right), \quad (42)$$

which is generated by the PI ((41), (28), and (29)), is an admissible control sequence for (1) on \mathbb{X} . The cost function $V^{(i)} = J(t_0)|_{\mathbf{u}=\mathbf{u}^{(i)}}$ satisfies the following Lyapunov equation (LE)

$$V_{\mathbf{x}_s}^{(i)T} [f_s(\mathbf{x}_s) + g_s(\mathbf{x}_s) \mathbf{u}^{(i)}] + [W_s(\mathbf{x}_s) + M(\mathbf{u}^{(i)})] = 0, \quad (43)$$

and then $V^*(\mathbf{x}_s) \leq V^{(i+1)}(\mathbf{x}_s) \leq V^{(i)}(\mathbf{x}_s)$ for $\forall \mathbf{x}_s \in \mathbb{X}$. In addition, $V^{(i)} \rightarrow V^*$ and $\mathbf{u}^{(i)} \rightarrow \mathbf{u}^*$ uniformly on \mathbb{X} .

Proof. The proof of this theorem follows the same lines of reasoning as in the proof of Lemma 1 and Theorem 1 in [38] and is omitted here for brevity. \square

Assumption 1. There exist $L_0 > 0$ and $\delta > 0$, such that for all $L > L_0$,

$$\frac{1}{L} \sum_{k=1}^L \theta_{i,k}^T \theta_{i,k} \geq I_{N_1+N_2}, \quad (44)$$

where $\theta_{i,k}$ is the k th row of $\Theta(\hat{k}^{(i)})$.

Assumption 2. The closed-loop system is composed of (18), and

$$\mathbf{u} = \mathbf{u}_0 + \mathbf{e} \quad (45)$$

is ISS (input-to-state stable) [52] when the exploration noise \mathbf{e} is considered as the input.

4. Case Study

In this section, two cases, including a CSTR (continuous stirred-tank reactor) model [14], are selected to test the feasibility and performance of the proposed control design.

4.1. *A Nonlinear System.* Consider the following nonlinear system:

$$\begin{aligned}\dot{x}_1 &= x_2, \\ \dot{x}_2 &= (1 - x_1)^2 x_2 - x_1 + u,\end{aligned}\quad (46)$$

with state and control constraints $x_2 \geq -0.4$ and $-1 \leq u \leq 1$, and initial condition $(x_1, x_2) = [1, 0]$.

To start with, one slack function was introduced:

$$-x_2 - 0.4 + \gamma e^{\alpha(t)} = 0. \quad (47)$$

According to (11) and (12), $\gamma = 0.4$ and $\alpha(0) = 0$. Denote $x_3 = \alpha$, and combine \dot{x}_3 with the original system:

$$\begin{aligned}\dot{x}_1 &= x_2, \\ \dot{x}_2 &= (1 - x_1)^2 x_2 - x_1 + u, \\ \dot{x}_3 &= \frac{1}{0.4e^{x_3}} [(1 - x_1)^2 x_2 - x_1 + u].\end{aligned}\quad (48)$$

For simulation purpose, we set $\mathbf{Q}_s = \text{diag}(1, 1, 0)$, so the augmented state x_3 is not taken into account in the value function, $\mathbf{R} = 0.5$; so, the value function mainly concerns the deviation between the actual state and its setting point, $\mathbf{H}(v) = \tanh(v)$. The basis functions for approximation of value function and control input were selected as polynomials of system states:

$$\begin{aligned}\Phi &= [x_1^2 \ x_2^2 \ x_1 x_2 \ x_1^4 \ x_2^4 \ x_1^2 x_2^2 \ x_1 x_2^3 \ x_1^3 x_2], \\ \Psi &= [x_1 \ x_2 \ x_1 x_2 \ x_1 x_2^2 \ x_1^2 x_2].\end{aligned}\quad (49)$$

The initial weight $\hat{k}^{(0)} = [0.8 \ -1.5 \ 2 \ 0 \ -1]$. The initial weight was determined using the structural information of system model and the basis functions. The actual input was set such that $\mathbf{u} = -\mathbf{H}(0.5\mathbf{R}^{-1}\hat{k}^{(0)}\Psi + 0.008e)$ with the exploration noise

$$\begin{aligned}e &= \text{rand} \cdot [\text{rand} \cdot \sin(0.1t) + \text{rand} \cdot \sin(t) + \text{rand} \cdot \sin(5t) \\ &\quad + \text{rand} \cdot \sin(10t)].\end{aligned}\quad (50)$$

The ADP iteration started after the learning finished at $t = 20$ s. After 20 iterations, the difference between $\hat{k}^{(20)}$ and $\hat{k}^{(19)}$ was less than 10^{-7} , and the ADP algorithm stopped. The resultant approximated optimal controller was $\hat{k}^{(20)} = [0.4586 \ 3.3914 \ 18.3107 \ 19.5615 \ 7.1479]$.

The state trajectories driven by the initial admissible controller and the approximated optimal controller were given in Figure 1. The approximated cost difference was illustrated in Figure 2. The control input using the two controllers were shown in Figure 3. From the figures, both the initial admissible controller and the approximated optimal controller can force the system to the equilibrium point $x_1 = 0$ and $x_2 = 0$. However, the state trajectories of x_1 and x_2 driven by the approximated optimal control has less fluctuations and less approximated control cost compared with the initial admissible controller.

4.2. *A CSTR Model.* Consider a reversible reaction $2A \rightleftharpoons B$ taken place in a CSTR, and the forward and reverse kinetic

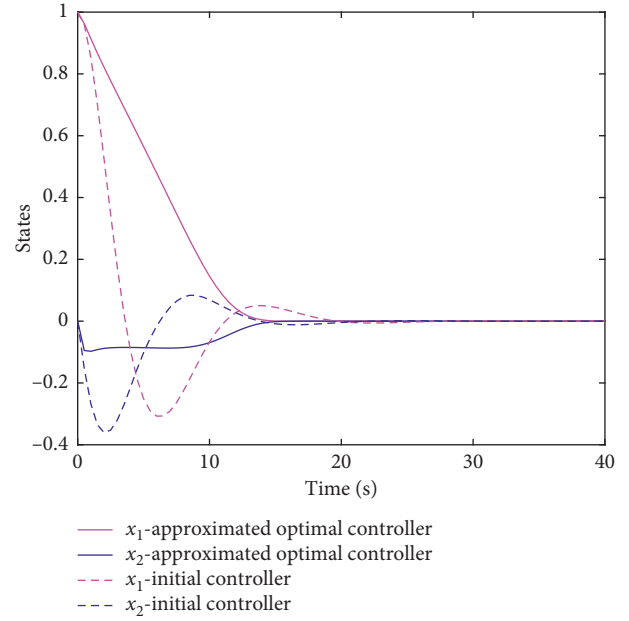


FIGURE 1: State trajectories driven by initial admissible control and approximated optimal control.

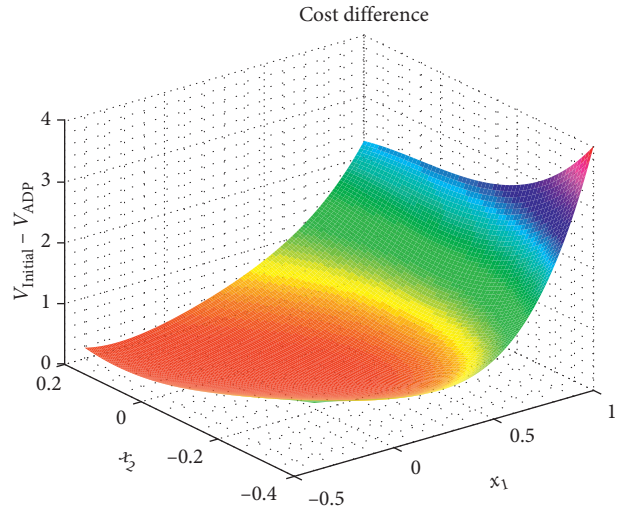


FIGURE 2: Cost difference between initial admissible control and approximated optimal control.

constants with respect to species B are k_a and k_b , respectively. The dynamics of the process is

$$\begin{aligned}\dot{x}_1 &= -2k_a x_1^2 - c_1 x_1 + 2k_b x_2 + \frac{F}{V} u, \\ \dot{x}_2 &= k_a x_1^2 + c_2 x_1 - \left(\frac{F}{V} + k_b\right) x_2,\end{aligned}\quad (51)$$

where $x_1 = C_A - \bar{C}_A$, $x_2 = C_B - \bar{C}_B$, and $u = C_{Ain} - \bar{C}_{Ain}$ are the deviation of the outlet concentrations of species A and B and the inlet concentration of species A , respectively. $c_1 = 4k_a \bar{C}_A + F/V$ and $c_2 = 2k_a \bar{C}_A$ are some constant variables. V and F are the volume and inlet flow-rate of the reactor,

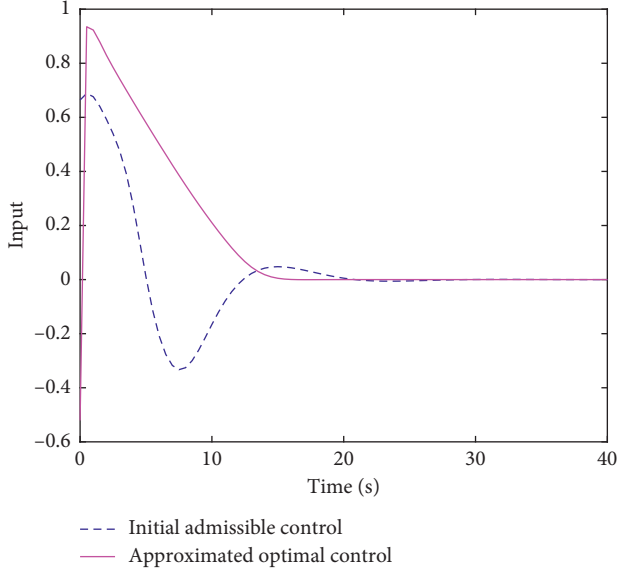


FIGURE 3: Control inputs using initial admissible controller and approximated optimal controller.

TABLE 1: Model parameters, initial conditions, and constraints.

Variable	Value
k_a	0.05
k_b	0.01
c_1	0.338
c_2	0.159
F/V	0.02
\bar{C}_A	1.59
\bar{C}_B	4.21
C_{Ain}	10
$x_1(0)$	0
$x_2(0)$	-1.71
$x_1(t)$	$[-0.09, 0.16]$
$x_2(t)$	≥ -4.21
$u(t)$	$[-10, 10]$

respectively. The constraints, initial conditions, and the values of model parameters were shown in Table 1.

According to Section 3.1, two slack functions were introduced first to transform the inequality constraints into equality constraints:

$$(x_1 + 0.09)(x_1 - 0.16) + \gamma_1 e^{\alpha_1(t)} = 0, \quad (52)$$

$$(-x_2 - 4.21) + \gamma_2 e^{\alpha_2(t)} = 0. \quad (53)$$

The values of γ_1 and γ_2 were calculated by setting both x_1 , α_1 and x_2 , α_2 as 0: $\gamma_1 = 0.0144$ and $\gamma_2 = 4.21$. The initial value of α_1 and α_2 were calculated using the initial value of x_1 and x_2 : $\alpha_1(0) = 0$ and $\alpha_2(0) = -0.5212$. By differentiating (51) and (52), the derivatives of α_1 and α_2 were obtained:

$$\dot{\alpha}_1 = \frac{2x_1 - 0.07}{0.0144e^{\alpha_1}} \left(-2k_a x_1^2 - c_1 x_1 + 2k_b x_2 + \frac{F}{V} u \right), \quad (54)$$

$$\dot{\alpha}_2 = \frac{1}{4.21e^{\alpha_2}} \left[k_a x_1^2 + c_2 x_1 - \left(\frac{F}{V} + k_b \right) x_2 \right]. \quad (55)$$

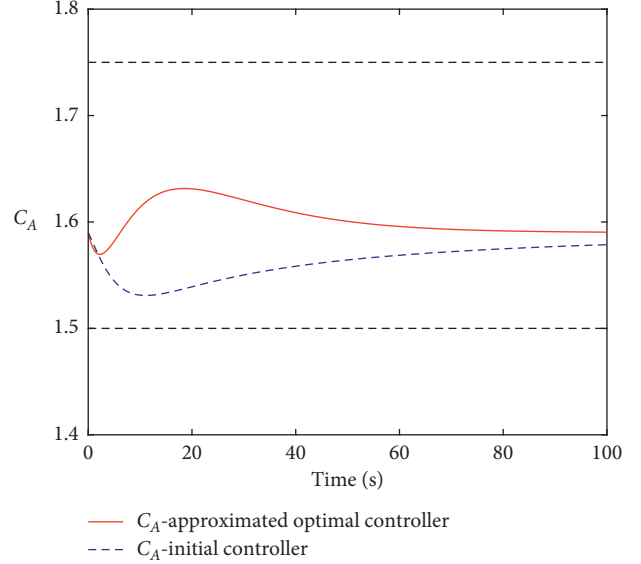


FIGURE 4: Trajectory of C_A driven by initial admissible control and approximated optimal control.

Then, (54) was differentiated until the appearance of input u which is present in \dot{x}_1 :

$$\ddot{\alpha}_2 = -\dot{\alpha}_2^2 + \frac{1}{4.21e^{\alpha_2}} [(0.1x_1 + 0.159)\dot{x}_1 - 0.03\dot{x}_2], \quad (56)$$

and $\dot{\alpha}_2(0) = 0.0205$.

Denote α_1 , α_2 , and $\dot{\alpha}_2$ as x_3 , x_4 , and x_5 , and combining the original system with the derived equality constraints, the following augmented system was obtained:

$$\begin{aligned} \dot{x}_1 &= -2k_a x_1^2 - c_1 x_1 + 2k_b x_2 + \frac{F}{V} u, \\ \dot{x}_2 &= k_a x_1^2 + c_2 x_1 - \left(\frac{F}{V} + k_b \right) x_2, \\ \dot{x}_3 &= \frac{2x_1 - 0.07}{0.0144e^{x_3}} \left(-2k_a x_1^2 - c_1 x_1 + 2k_b x_2 + \frac{F}{V} u \right), \\ \dot{x}_4 &= x_5, \\ \dot{x}_5 &= -x_4^2 + \frac{1}{4.21e^{x_4}} \left\{ (0.1x_1 + 0.159) \right. \\ &\quad \cdot \left(-2k_a x_1^2 - c_1 x_1 + 2k_b x_2 + \frac{F}{V} u \right) \\ &\quad \left. - 0.03 \left[k_a x_1^2 + c_2 x_1 - \left(\frac{F}{V} + k_b \right) x_2 \right] \right\}. \end{aligned} \quad (57)$$

In the simulation, $\mathbf{Q}_s = \text{diag}(5, 5, 0, 0, 0)$; so, the augmented states were not considered when evaluating the control performance. Compared with the evolution of outlet concentrations to their set points in time, the selection of input variables is not that important, so $\mathbf{R} = 0.5$ and $\mathbf{H}(v) = 10 \tanh(v/10)$. The basis functions in Φ include x_1^2 , x_2^2 , $x_1 x_2$, x_1^4 , x_2^4 , $x_1^2 x_2^2$, $x_1 x_2^3$, $x_1^3 x_2$, x_1^6 , x_2^6 , $x_1 x_2^5$, $x_1^2 x_2^4$, $x_1^3 x_2^3$, $x_1^4 x_2^2$, $x_1^5 x_2$, and $\Psi = [x_1 \ x_2 \ x_1^2 \ x_2^2]$. The initial weight $\hat{k}^{(0)} = [-1 \ -1 \ -1 \ -1]$. The exploration noise was set such

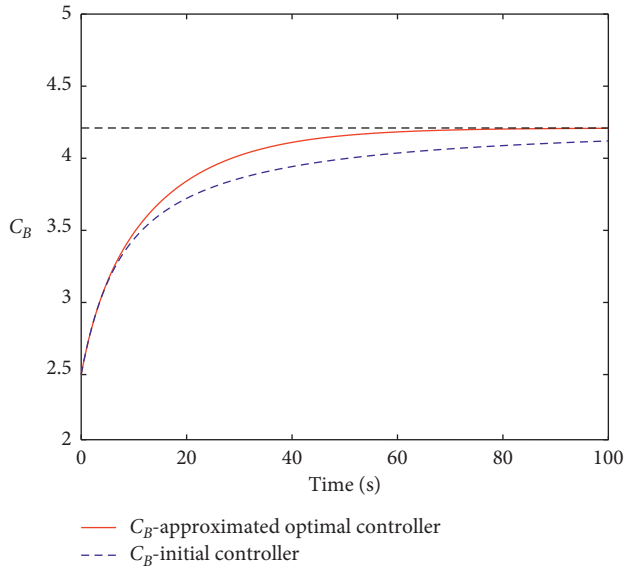


FIGURE 5: Trajectory of C_B driven by initial admissible control and approximated optimal control.

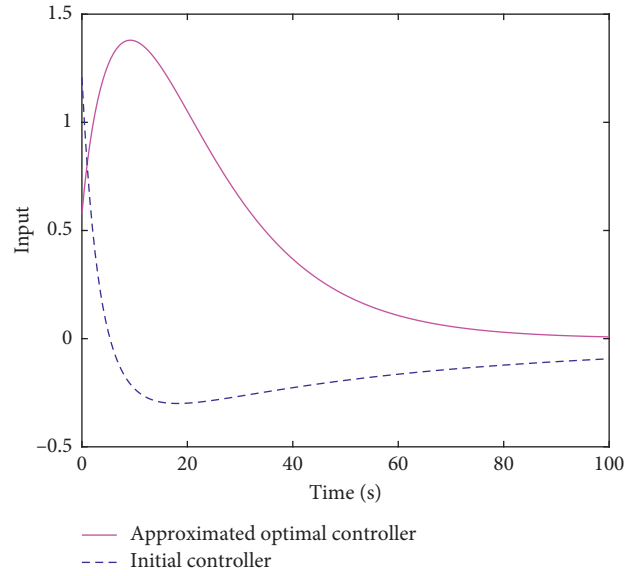


FIGURE 7: Control inputs using initial admissible controller and approximated optimal controller.

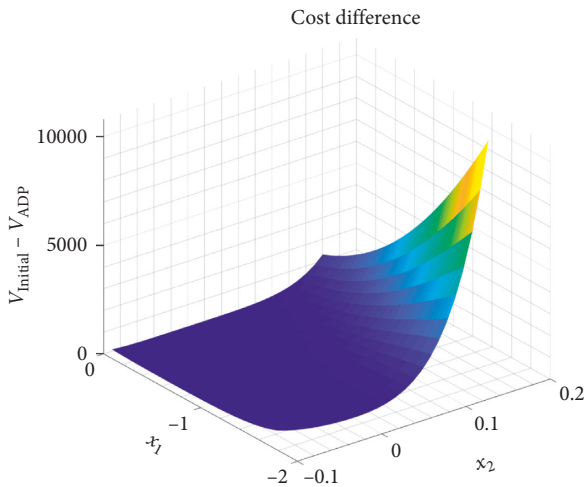


FIGURE 6: Cost difference between initial admissible control and approximated optimal control.

that $\mathbf{u} = -\mathbf{H}(0.5\mathbf{R}^{-1}\hat{\mathbf{k}}^{(0)}\Psi + 0.1\mathbf{e})$ with $\mathbf{e} = \text{rand} \cdot [\text{rand} \cdot \sin(0.1t) + \text{rand} \cdot \sin(t) + \text{rand} \cdot \sin(5t) + \text{rand} \cdot \sin(10t)]$.

The ADP iteration started after the learning finished at $t = 40$ s. After 20 iterations, the difference between $\hat{\mathbf{k}}^{(20)}$ and $\hat{\mathbf{k}}^{(19)}$ was less than 10^{-7} , and the ADP algorithm stopped. The resultant approximated optimal controller was $\hat{\mathbf{k}}^{(20)} = [-6.1373 \ 2.7789 \ -3.2045 \ 1.4604]$.

The state trajectories of x_1 and x_2 under the initial admissible controller and the approximated optimal controller were given in Figures 4 and 5, respectively. The approximated cost difference was illustrated in Figure 6. The control input using the two controllers were shown in Figure 7. Compared with the initial admissible controller,

the approximated optimal controller could force the CSTR to the desired working point $C_A = 1.59$ and $C_B = 4.21$ with less steady-state error and better performance regarding to the performance index.

In the simulation, structural knowledge of the system model was required to derive the explicit formulations of (47) and (56). However, the numerical value of the slack variables can be obtained without knowing the full system dynamics; only the model structure is required. Thus, via learning from online input and state information, the performance of the controller was improved.

5. Conclusions

An approximated optimal control approach for chemical processes with both input and state constraints was proposed. Its feasibility and performance were tested via two nonlinear examples. The proposed approach requires only the structure knowledge of the system model. The value information of the model parameters is not needed. This indicates the proposed approach can be applied to practical systems or a single unit (like a CSTR in this study) with determined model structure but has unknown parameters under some working conditions. However, the proposed approach is applicable only when Assumptions 1 and 2 hold. The global stability issue still needs future study.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The author declares that there are no conflicts of interest regarding the publication of this paper.

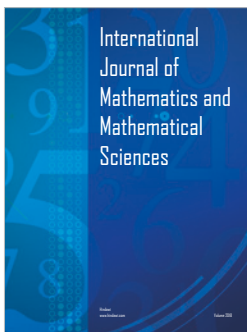
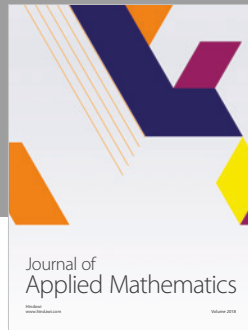
Acknowledgments

This work was supported by the National Natural Science Foundation of China (grant no. 61703441).

References

- [1] A. Saberi, Z.-L. Lin, and A. R. Teel, "Control of linear systems with saturating actuators," *IEEE Transactions on Automatic Control*, vol. 41, no. 3, pp. 368–378, 1996.
- [2] D. Q. Mayne, J. B. Rawlings, C. V. Rao, and P. O. M. Sokaert, "Constrained model predictive control: stability and optimality," *Automatica*, vol. 36, no. 6, pp. 789–814, 2000.
- [3] N. H. El-Farra and P. D. Christofides, "Integrating robustness, optimality and constraints in control of nonlinear processes," *Chemical Engineering Science*, vol. 56, no. 5, pp. 1841–1868, 2001.
- [4] R. M. Ansari and M. O. Tade, "Constrained nonlinear multivariable control of a catalytic reforming process," *Control Engineering Practice*, vol. 6, no. 6, pp. 695–706, 1998.
- [5] G. Andrikopoulos, G. Nikolakopoulos, I. Arvanitakis, and S. Manesis, "Piecewise affine modeling and constrained optimal control for a pneumatic artificial muscle," *IEEE Transactions on Industrial Electronics*, vol. 61, no. 2, pp. 904–916, 2014.
- [6] G. Leitmann, *Optimization Techniques: With Applications to Aerospace Systems*, Vol. 5, Academic Press, London, UK, 1962.
- [7] A. E. Bryson and W. F. Denham, "Optimal programming problems with inequality constraints. ii - solution by steepest-ascent," *AIAA Journal*, vol. 2, no. 1, pp. 25–34, 1964.
- [8] M. M. Lele and D. H. Jacobson, "A proof of the convergence of the kelley-bryson penalty function technique for state-constrained control problems," *Journal of Mathematical Analysis and Applications*, vol. 26, no. 1, pp. 163–169, 1969.
- [9] Y. Sakawa, "On a solution of an optimization problem in linear systems with quadratic performance index," *SIAM Journal on Control*, vol. 4, no. 2, pp. 382–395, 1966.
- [10] D. Jacobson and M. Lele, "A transformation technique for optimal control problems with a state variable inequality constraint," *IEEE Transactions on Automatic Control*, vol. 14, no. 5, pp. 457–464, 1969.
- [11] W. W. Hager, "The Ritz-Trefftz method for state and control constrained optimal control problems," *SIAM Journal on Numerical Analysis*, vol. 12, no. 6, pp. 854–867, 1975.
- [12] J. Vlassenbroeck, "A Chebyshev polynomial method for optimal control with state constraints," *Automatica*, vol. 24, no. 4, pp. 499–506, 1988.
- [13] A. E. B. Lim, J. B. Moore, and L. Faybusovich, "Separation theorem for linearly constrained LQG optimal control," *Systems & Control Letters*, vol. 28, no. 4, pp. 227–235, 1996.
- [14] V. Manousiouthakis and D. J. Chmielewski, "On constrained infinite-time nonlinear optimal control," *Chemical Engineering Science*, vol. 57, no. 1, pp. 105–114, 2002.
- [15] A. Balestrino, A. Caiti, and S. Grammatico, "Multivariable constrained process control via lyapunov R-functions," *Journal of Process Control*, vol. 22, no. 9, pp. 1762–1772, 2012.
- [16] T. Kiefer, K. Graichen, and A. Kugi, "Trajectory tracking of a 3DOF laboratory helicopter under input and state constraints," *IEEE Transactions on Control Systems Technology*, vol. 18, no. 4, pp. 944–952, 2010.
- [17] G. Stathopoulos, M. Korda, and C. N. Jones, "Solving the infinite-horizon constrained lqr problem using accelerated dual proximal methods," *IEEE Transactions on Automatic Control*, vol. 62, no. 4, pp. 1752–1767, 2017.
- [18] W. C. Li and L. T. Biegler, "Process control strategies for constrained nonlinear systems," *Industrial & Engineering Chemistry Research*, vol. 27, no. 8, pp. 1421–1433, 1988.
- [19] B. J. Allison and J. B. Ball, "Constrained model predictive control of blow tank consistency," *Control Engineering Practice*, vol. 12, no. 7, pp. 837–845, 2004.
- [20] X.-S. Chen, Q. Li, and S.-M. Fei, "Constrained model predictive control in ball mill grinding process," *Powder Technology*, vol. 186, no. 1, pp. 31–39, 2008.
- [21] S. Dubljevic, "Model predictive control of Kuramoto-Sivashinsky equation with state and input constraints," *Chemical Engineering Science*, vol. 65, no. 15, pp. 4388–4396, 2010.
- [22] L. Lao, M. Ellis, and P. D. Christofides, "Handling state constraints and economics in feedback control of transport-reaction processes," *Journal of Process Control*, vol. 32, pp. 98–108, 2015.
- [23] J. L. Godoy, A. H. González, and J. E. Normey-Rico, "Constrained latent variable model predictive control for trajectory tracking and economic optimization in batch processes," *Journal of Process Control*, vol. 45, pp. 1–11, 2016.
- [24] L. Roca, J. L. Guzman, J. E. Normey-Rico, M. Berenguel, and L. Yebra, "Robust constrained predictive feedback linearization controller in a solar desalination plant collector field," *Control Engineering Practice*, vol. 17, no. 9, pp. 1076–1088, 2009.
- [25] P. M. Marusak and S. Kuntanapreeda, "Constrained model predictive force control of an electrohydraulic actuator," *Control Engineering Practice*, vol. 19, no. 1, pp. 62–73, 2011.
- [26] J. M. Grosso, C. Ocampo-Martínez, V. Puig, and B. Joseph, "Chance-constrained model predictive control for drinking water networks," *Journal of Process Control*, vol. 24, no. 5, pp. 504–516, 2014.
- [27] P. O. M. Sokaert and J. B. Rawlings, "Constrained linear quadratic regulation," *IEEE Transactions on Automatic Control*, vol. 43, no. 8, pp. 1163–1169, 1998.
- [28] L.-X. Wang and F. Wan, "Structured neural networks for constrained model predictive control," *Automatica*, vol. 37, no. 8, pp. 1235–1243, 2001.
- [29] A. Bemporad, M. Morari, V. Dua, and E. N. Pistikopoulos, "The explicit linear quadratic regulator for constrained systems," *Automatica*, vol. 38, no. 1, pp. 3–20, 2002.
- [30] G. Pin, M. Filippo, F. A. Pellegrino, G. Fenu, and T. Parisini, "Approximate model predictive control laws for constrained nonlinear discrete-time systems: analysis and offline design," *International Journal of Control*, vol. 86, no. 5, pp. 804–820, 2013.
- [31] P. Mhaskar, N. H. El-Farra, and P. D. Christofides, "Stabilization of nonlinear systems with state and control constraints using lyapunov-based predictive control," *Systems & Control Letters*, vol. 55, no. 8, pp. 650–659, 2006.
- [32] E. Polak, "An historical survey of computational methods in optimal control," *SIAM Review*, vol. 15, no. 2, pp. 553–584, 1973.
- [33] R. F. Hartl, S. P. Sethi, and R. G. Vickson, "A survey of the maximum principles for optimal control problems with state constraints," *SIAM Review*, vol. 37, no. 2, pp. 181–218, 1995.
- [34] D. Q. Mayne, "Model predictive control: recent developments and future promise," *Automatica*, vol. 50, no. 12, pp. 2967–2986, 2014.
- [35] S. Yin, H. Gao, and O. Kaynak, "Data-driven control and process monitoring for industrial applications—part I," *IEEE*

- Transactions on Industrial Electronics*, vol. 61, no. 11, pp. 6356–6359, 2014.
- [36] A. Chakrabarty, V. Dinh, M. J. Corless, A. E. Rundell, S. H. Zak, and G. T. Buzzard, “Support vector machine informed explicit nonlinear model predictive control using low-discrepancy sequences,” *IEEE Transactions on Automatic Control*, vol. 62, no. 1, pp. 135–148, 2017.
- [37] W.-S. Lin and C.-H. Zheng, “Constrained adaptive optimal control using a reinforcement learning agent,” *Automatica*, vol. 48, no. 10, pp. 2614–2619, 2012.
- [38] M. Abu-Khalaf and F. L. Lewis, “Nearly optimal control laws for nonlinear systems with saturating actuators using a neural network HJB approach,” *Automatica*, vol. 41, no. 5, pp. 779–791, 2005.
- [39] H. Modares, F. L. Lewis, and M.-B. Naghibi-Sistani, “Adaptive optimal control of unknown constrained-input systems using policy iteration and neural networks,” *IEEE Transactions on Neural Networks and Learning Systems*, vol. 24, no. 10, pp. 1513–1525, 2013.
- [40] B. Luo, H.-N. Wu, T. Huang, and D. Liu, “Reinforcement learning solution for hjb equation arising in constrained optimal control problem,” *Neural Networks*, vol. 71, pp. 150–158, 2015.
- [41] X. Yang, D. Liu, B. Luo, and C. Li, “Data-based robust adaptive control for a class of unknown nonlinear constrained-input systems via integral reinforcement learning,” *Information Sciences*, vol. 369, pp. 731–747, 2016.
- [42] Q. Zhang, D. Zhao, and Y. Zhu, “Data-driven adaptive dynamic programming for continuous-time fully cooperative games with partially constrained inputs,” *Neurocomputing*, vol. 238, pp. 377–386, 2017.
- [43] Y. Zhu, D. Zhao, H. He, and J. Ji, “Event-triggered optimal control for partially unknown constrained-input systems via adaptive dynamic programming,” *IEEE Transactions on Industrial Electronics*, vol. 64, no. 5, pp. 4101–4109, 2017.
- [44] F. L. Lewis and D. Vrabie, “Reinforcement learning and adaptive dynamic programming for feedback control,” *IEEE Circuits and Systems Magazine*, vol. 9, no. 3, pp. 32–50, 2009.
- [45] W. Gao and Z.-P. Jiang, “Adaptive dynamic programming and adaptive optimal output regulation of linear systems,” *IEEE Transactions on Automatic Control*, vol. 61, no. 12, pp. 4164–4169, 2016.
- [46] Q. Wei, D. Liu, F. L. Lewis, Y. Liu, and J. Zhang, “Mixed iterative adaptive dynamic programming for optimal battery energy control in smart residential microgrids,” *IEEE Transactions on Industrial Electronics*, vol. 64, no. 5, pp. 4110–4120, 2017.
- [47] Q.-Y. Fan and G.-H. Yang, “Nearly optimal sliding mode fault-tolerant control for affine nonlinear systems with state constraints,” *Neurocomputing*, vol. 216, pp. 78–88, 2016.
- [48] R. Chi, X. Liu, R. Zhang, Z. Hou, and B. Huang, “Constrained data-driven optimal iterative learning control,” *Journal of Process Control*, vol. 55, pp. 10–29, 2017.
- [49] G. N. Saridis and C.-S. G. Lee, “An approximation theory of optimal control for trainable manipulators,” *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 9, no. 3, pp. 152–159, 1979.
- [50] Y. Jiang and Z.-P. Jiang, “Robust adaptive dynamic programming and feedback stabilization of nonlinear systems,” *IEEE Transactions on Neural Networks and Learning Systems*, vol. 25, no. 5, pp. 882–893, 2014.
- [51] S. Lyashevskiy, “Constrained optimization and control of nonlinear systems: new results in optimal control,” in *Proceedings of the 35th IEEE Conference on Decision and Control*, pp. 541–546, IEEE, Kobe, Japan, December 1996.
- [52] H. K. Khalil, *Nonlinear Systems*, Prentice-Hall, New Jersey, NJ, USA, 3rd edition, 2002.



Hindawi

Submit your manuscripts at
www.hindawi.com

