Computers and Chemical Engineering 153 (2021) 107411

Contents lists available at ScienceDirect

Computers and Chemical Engineering

journal homepage: www.elsevier.com/locate/compchemeng

Sparse-identification-based model predictive control of nonlinear two-time-scale processes

Fahim Abdullah[a], Zhe Wu[a], Panagiotis D. Christofides[a,b,*]

[a] Department of Chemical and Biomolecular Engineering, University of California, Los Angeles, CA, 90095-1592, USA
[b] Department of Electrical and Computer Engineering, University of California, Los Angeles, CA 90095-1592, USA

ARTICLE INFO

Article history:
Received 10 April 2021
Revised 11 June 2021
Accepted 13 June 2021
Available online 17 June 2021

Keywords:
Two-time-scale processes
Nonlinear processes
Singular perturbations
Model predictive control
Sparse identification
Chemical processes

ABSTRACT

This paper focuses on the design of model predictive controllers for nonlinear two-time-scale processes using only process measurement data. By first identifying and isolating the slow and fast variables in a two-time-scale process, the model predictive controller is designed based on the reduced slow subsystem consisting of only the slow variables, since the fast states can deteriorate controller performance when directly included in the model used in the controller. In contrast to earlier works, in the present work, the reduced slow subsystem is constructed from process data using sparse identification, which identifies nonlinear dynamical systems as first-order ordinary differential equations using an efficient, convex algorithm that is highly optimized and scalable. Results from the mathematical framework of singular perturbations are combined with standard assumptions to derive sufficient conditions for closed-loop stability of the full singularly perturbed closed-loop system. The effectiveness of the proposed controller design is illustrated via its application to a non-isothermal reactor with the concentration and temperature profiles evolving in different time-scales, where it is found that the controller based on the sparse identified slow subsystem can achieve superior closed-loop performance *versus* existing approaches for the same controller parameters.

1. Introduction

Chemical processes such as fluidized catalytic crackers, biochemical reactors and distillation columns, often involve physicochemical phenomena occurring in vastly-different time scales. In chemical process modeling, time-scale separation has been handled via singular perturbation techniques (*e.g.*, Kumar et al., 1998) in the context of first-principles modeling with the goal of building well-conditioned ordinary differential equation (ODE) models that can be used for process analysis and controller design. More recently, in Alanqar et al. (2015), a constraint was incorporated in the optimization problem of fitting a polynomial nonlinear state-space model to nonlinear process data to avoid identified model ill-conditioning (stiffness) in the construction of nonlinear ODE models. To control a two-time-scale system, a composite control system using multi-rate sampling was proposed in Chen et al. (2011), where a "fast" feedback controller was designed to stabilize the fast dynamics subsystem, and a model predictive controller (MPC) was developed to stabilize the slow dynamic and optimize closed-loop performance. Additionally, in Ellis et al. (2013), a composite controller which consisted of a Lyapunov-based MPC for the fast subsystem, and a Lyapunov-based Economic MPC for the slow subsystem, was developed to operate the system in a time-varying fashion in order to improve process economics while achieving desired closed-loop stability properties. In all the above efforts, an explicit nonlinear process model, typically obtained from first-principles, was assumed to be available.

Model predictive control (MPC) is an optimization-based process control scheme that can optimize process performance and account for state and control input constraints. Since the MPC optimization problem relies on a process model for predicting the future state evolution, the performance of MPC relies heavily on its model quality. In recent years, machine learning techniques have been applied to model chemical processes from data when first-principle models are unavailable. For example, in Wu et al. (2019a,b), recurrent neural networks were used to build the data-driven model for a general class of nonlinear processes, and were then incorporated in Lyapunov-based MPC to ensure desired closed-loop stability and performance. In Abdullah et al. (2021), reduced-order data-driven models were developed for nonlinear two-time-scale systems using nonlinear principal component analysis and sparse identification methods. The simultaneous design and control of nonlinear dynamical systems via the use of robust

* Corresponding author.
E-mail address: pdc@seas.ucla.edu (P.D. Christofides).

https://doi.org/10.1016/j.compchemeng.2021.107411
0098-1354/

tools based on Lyapunov theory has been investigated generally in Ricardez Sandoval et al. (2008); Sanchez-Sanchez and Ricardez-Sandoval (2013) and also under uncertainty in the context of MPC in Ricardez-Sandoval et al. (2009).

In the field of nonlinear dynamical modeling, the technique of sparse identification developed in Brunton et al. (2016b) has successfully been applied to a diverse array of nonlinear regression problems, including chaotic systems and nonlinear partial differential equations. However, its application to two-time-scale systems has only recently been investigated. In Champion et al. (2019), the authors extended the sparse identification procedure to the discovery of two-time-scale systems when full or even partial measurements of the variables are available under the assumption of linear coupling across the time-scales. In particular, when full state measurements were available, it was shown that, although sparse identification could correctly identify uni-time-scale systems with as little data as 5% of a period, for two-time-scale systems, the data requirement increased linearly with the ratio of slow to fast dynamics under naive uniform sampling. This was due to the fast trajectories requiring a smaller sampling period. In fact, even in uni-time-scale systems, when the aforementioned 5% of data was sufficient to rebuild the system, the data had to be captured at the relatively fast segments of the period at a very high sampling rate, as the information contained in the faster regions was much more than the information captured in the slower regions of the period. An alternative to overcome the linear increase in data requirement with diverging time-scale-multiplicities was proposed and devised in Champion et al. (2019) known as burst sampling, which allowed sparse identification to identify both the slow and fast subsystems with a significantly smaller fragment of the data over one period. However, since the present work aims to utilize sparse identification for the purpose of control, identifying the fast subsystem as a stiff differential equation via a more complicated composite multirate sampling is not advantageous. Instead, sparse identification is only used to identify the slow subsystem, which is incorporated into the controller. The application of sparse identification in closed-loop control is an area that has not been investigated in-depth yet.

Inspired by these results, in this work, we use time-series data from the slow process variables (subset of process state variables in some coordinate system) to construct well-conditioned data-driven models for a general class of two-time-scale nonlinear processes, and evaluate their performance in the context of MPC of two-time-scale processes. The rest of this article is organized as follows: in Section 2, the notations, and the class of nonlinear two-time-scale systems considered are given. In Section 3, the sparse identification (SI) model for the slow subsystem is introduced. In Section 4, the formulation of LMPC using SI models is presented, while the closed-loop singularly perturbed system stability analysis showing the boundedness of closed-loop states in a small neighborhood around the origin is given in Section 5. In Section 6, a chemical reactor example which exhibits two-time-scale behavior is simulated to demonstrate the effectiveness of the proposed modeling and control approaches.

2. Preliminaries

2.1. Notation

The notation $|\cdot|$ denotes the Euclidean norm of a vector. $x^\top$ is used to denote the transpose of x. The notation $L_f V(x)$ denotes the standard Lie derivative $L_f V(x) := \frac{\partial V(x)}{\partial x} f(x)$. Set subtraction is denoted by "\", *i.e.*, $A \backslash B := \{x \in \mathbb{R}^n \mid x \in A, x \notin B\}$. The function $f(\cdot)$ is of class $\mathcal{C}^1$ if it is continuously differentiable in its domain. A continuous function $\alpha : [0, a) \to [0, \infty)$ is said to belong to class $\mathcal{K}$ if it is strictly increasing and is zero only when evaluated at zero. A continuous function $\beta : [0, a) \times [0, \infty) \to [0, \infty)$ is said to belong to class $\mathcal{KL}$ if, for each fixed t, the function $\beta(\cdot, t)$ is of class $\mathcal{K}$, while, for each fixed s, the function $\beta(s, \cdot)$ is decreasing and tends to zero as $s \to \infty$.

2.2. Class of systems

The general class of two-time-scale continuous-time nonlinear systems with m states considered in this work has the following general form:

$$\dot{x} = f_1(x, z, u, \epsilon) \tag{1a}$$

$$\epsilon \dot{z} = f_2(x, z, \epsilon) \tag{1b}$$

where $x \in \mathbb{R}^n$ and $z \in \mathbb{R}^p$ are the slow and fast state vectors, respectively, with $n + p = m$. $u \in \mathbb{R}^q$ is the manipulated input vector with constraints defined by $u \in U := \{u_i^{\min} \le u_i \le u_i^{\max}, i = 1, \ldots, q\} \subset \mathbb{R}^q$. ϵ is a small positive parameter representing the speed ratio of the slow to the fast dynamics of the system. The vector functions $f_1(x, z, u, \epsilon)$ and $f_2(x, z, \epsilon)$ are sufficiently smooth vector functions in $\mathbb{R}^n$ and $\mathbb{R}^p$, respectively. In the system of Eq. (1), after a short transient period, the fast states, z, converge to a slow manifold (provided such an isolated manifold exists) and can be expressed by a nonlinear algebraic expression in x, the slow states. Therefore, following the standard two-time scale decomposition procedure in Kokotović et al. (1986), we can set $\epsilon = 0$ in Eq. (1) and obtain:

$$\dot{x} = f_1(x, z, u, 0) \tag{2a}$$

$$0 = f_2(x, z, 0) \tag{2b}$$

Remark 1. The class of singularly perturbed system of Eq. 1 are presented for analysis purposes, and will not be used to derive models that will be used in the controllers. Furthermore, to simplify the development, we focus on two-time-scale processes with stable fast dynamic, and for this reason, we take $f_2(x, z, \epsilon)$ to be independent of u. However, our analysis can be readily adapted to deal with systems in which $f_2(\cdot)$ is a function of u.

Assumption 1. Eq. (2b) possesses a unique root given by

$$z_s = \hat{f}_2(x) \tag{3}$$

where z_s is a quasi-steady state for the fast state z, and $\hat{f}_2 : \mathbb{R}^n \to \mathbb{R}^p$ and its derivative are locally Lipschitz continuous.

Assumption 1 is standard in the singular perturbation framework as it ensures that the system has an isolated equilibrium manifold for the fast dynamics on which z can be written as an algebraic function of x. Substituting Eq. (3) into Eq. (2a) gives the reduced slow subsystem,

$$\dot{x} = f_1(x, \hat{f}_2(x), u, 0) \tag{4}$$

For the fast state, we define a fast time scale $\tau = t/\epsilon$ and a new coordinate $y := z - \hat{f}_2(x)$. Rewriting Eq. (1b) as a derivative with respect to τ rather than t and setting $\epsilon = 0$ yields the fast subsystem,

$$\frac{dy}{d\tau} = f_2(x, \hat{f}_2(x) + y, 0) \tag{5}$$

We assume the input appears linearly in Eq. (4) and, therefore, rewrite the slow subsystem of Eq. (4) in the following form throughout the manuscript:

$$\dot{x} = F(x, u) := f(x) + g(x)u, \qquad x(t_0) = x_0 \tag{6}$$

where $f(\cdot)$ and $g(\cdot)$ are sufficiently smooth vector and matrix functions of dimensions $n \times 1$ and $n \times q$, respectively. Without loss of

generality, throughout the manuscript, the initial time t_0 is taken to be zero ($t_0 = 0$), and it is assumed that $f(0) = 0$, and thus, the origin is a steady-state of the nonlinear system of Eq. (6), *i.e.*, $(x_s^*, u_s^*) = (0, 0)$, where x_s^* and u_s^* represent the steady-state slow state and input vectors, respectively.

Assumption 2. The origin of the closed-loop fast subsystem of Eq. (5) is globally asymptotically stable, uniformly in x in the sense that there exists a function β_y of class $\mathcal{KL}$ such that for any $y(0) \in \mathbb{R}^p$,

$$|y(t)| \leq \beta_y\left(|y(0)|, \frac{t}{\epsilon}\right) \qquad \forall\, t \geq 0 \tag{7}$$

2.3. Stabilizability assumption via control Lyapunov function

With respect to the stabilizability requirement for the slow dynamics, it is assumed that there exists a stabilizing control law $u = \Phi(x) \in U$ (*e.g.*, the universal Sontag control law (Lin and Sontag, 1991)) such that the origin of the nominal slow subsystem of Eq. (6) is rendered asymptotically stable in an open neighborhood D around the origin in the sense that there exist a $\mathcal{C}^1$ control Lyapunov function, $V(x)$, and four functions, a_1, a_2, a_3, a_4 of class $\mathcal{K}$ such that $\forall\, x \in \mathbb{R}^n$:

$$a_1(|x|) \leq V(x) \leq a_2(|x|), \tag{8a}$$

$$\dot{V}(x) = \frac{\partial V(x)}{\partial x} F(x, \Phi(x)) \leq -a_3(|x|), \tag{8b}$$

$$\left|\frac{\partial V(x)}{\partial x}\right| \leq a_4(|x|) \tag{8c}$$

The Sontag law, a candidate controller for $\Phi(x)$, is given in the following form:

$$\varphi_i(x) = \begin{cases} -\dfrac{p + \sqrt{p^2 + q^4}}{q^\top q} q & \text{if} \quad q \neq 0 \\ 0 & \text{if} \quad q = 0 \end{cases} \tag{9a}$$

$$\Phi_i(x) = \begin{cases} u_i^{\min} & \text{if} \quad \varphi_i(x) < u_i^{\min} \\ \varphi_i(x) & \text{if} \quad u_i^{\min} \leq \varphi_i(x) \leq u_i^{\max} \\ u_i^{\max} & \text{if} \quad \varphi_i(x) > u_i^{\max} \end{cases} \tag{9b}$$

where p denotes $L_f V(x)$ and q denotes $L_{g_i} V(x)$, $f = [f_1 \cdots f_n]^\top$, $g_i = [g_{1i}, \ldots, g_{ni}]^\top$, $i = 1, 2, \ldots, q$. $\varphi_i(x)$ of Eq. (9a) represents the i_{th} component of the control law $\varphi(x)$. $\Phi_i(x)$ of Eq. (9) represents the i_{th} component of the saturated control law $\Phi(x)$ that accounts for the input constraint $u \in U$.

First, a region where the conditions of Eq. (8) are satisfied under the controller $u = \Phi(x) \in U$ as $\phi_u = \{x \in \mathbb{R}^n \mid \dot{V}(x) = L_f V + L_g V u < -kV(x), u = \Phi(x) \in U\} \cup \{0\}$, where $k > 0$, is characterized. Then the closed-loop stability region Ω_ρ for the nonlinear slow subsystem of Eq. (6) is defined as a level set of the Lyapunov function, which is inside ϕ_u: $\Omega_\rho := \{x \in \phi_u \mid V(x) \leq \rho\}$, where $\rho > 0$ and $\Omega_\rho \subset \phi_u$. Furthermore, the Lipschitz property of $F(x, u)$ combined with the bound on u implies that there exist positive constants M, L, L' such that the following inequalities hold $\forall\, x, x' \in D$, $\forall\, u \in U$:

$$|F(x, u)| \leq M \tag{10a}$$

$$|F(x, u) - F(x', u)| \leq L|x - x'| \tag{10b}$$

$$\left|\frac{\partial V(x)}{\partial x} F(x, u) - \frac{\partial V(x')}{\partial x} F(x', u)\right| \leq L'|x - x'| \tag{10c}$$

3. Sparse identification model for the slow subsystem

3.1. Sparse identification

Sparse identification is a nonlinear system identification method used to model dynamical systems. The application of sparsity methods to dynamic system modeling can be found in the recent literature (Wang et al., 2011; Schaeffer et al., 2013; Ozolinš et al., 2013; Mackey et al., 2014; Brunton et al., 2014; Proctor et al., 2014; Bai et al., 2015; Arnaldo et al., 2015). Sparse identification is used to approximately reconstruct a continuous-time ODE of the form,

$$\dot{\hat{x}} = \hat{f}(\hat{x}) \tag{11}$$

using only numerical data from the system without requiring knowledge of the underlying physics of the process.

Sparse identification exploits the sparsity of the right-hand side of Eq. (11) since $\hat{f}(\hat{x})$ contains very few nonzero terms in practical systems, rendering it sparse in a higher-dimensional space of candidate nonlinear functions. The nonzero terms can then be calculated using scalable convex methods. To carry out sparse identification, the open-loop process is simulated over a wide range of initial conditions. From the resulting data, the slow state measurements are sampled with a sufficiently small sampling period and concatenated into a data matrix, X, of the form,

$$X = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix} \tag{12}$$

where each x_i is a column of time-series data for state i for $i = 1, \ldots, n$. The time-derivative of X, denoted by $\dot{X}$, is estimated in this work using second-order central finite differences since the noise-free case is considered and since the sampling period is sufficiently small. Subsequently, a function library, $\Theta(X)$, is created consisting of r nonlinear functions of the columns of X. The r functions represent possible terms for f, the right-hand side of Eq. (6). The goal of the sparse identification algorithm is to identify the active terms in this library by taking advantage of sparsity. The augmented library, $\Theta(X)$, considered in this work is of the form,

$$\Theta(X) = \begin{bmatrix} | & | & | & | & | & | & | & | \\ \mathbf{1} & X & X^{P_2} & X^{P_3} & \sin X & \cos X & \tan X & \tanh X \\ | & | & | & | & | & | & | & | \end{bmatrix} \tag{13}$$

where, for example, X^{P_2} denotes all quadratic nonlinearities, given by

$$X^{P_2} = \begin{bmatrix} x_1^2 & x_1 x_2 & \cdots & x_2^2 & x_2 x_3 & \cdots & x_n^2 \end{bmatrix} \tag{14}$$

The above choice of candidate nonlinear functions was motivated by the fact that polynomials and trigonometric functions form a basis for many practical systems as mentioned in Brunton et al. (2016b).

Remark 2. As this is a novel method, only the noise-free case is considered in this work. However, if the method is applied to noisy data, the time-derivative $\dot{X}$ cannot be computed using regular finite difference methods as such methods will amplify the noise present and not yield meaningful values. Instead, the derivative approximation will require the application of more advanced techniques such as the total-variation regularized derivative (Chartrand, 2011) or smoothed finite difference, where the noisy data is first smoothed out using a filter (such as the Savitzky-Golay filter) before calculating the finite differences (Savitzky and Golay, 1964).

In sparse identification, for each of the n slow states, we determine the r coefficients that pre-multiply the r candidate nonlinear functions considered in the function library, $\Theta(X)$. Denoting each

corresponding coefficient vector by ξ, the n coefficient vectors can be compactly written in matrix form as

$$\Xi = \begin{bmatrix} \xi_1 & \xi_2 & \cdots & \xi_n \end{bmatrix} \tag{15}$$

where each $\xi_i \in \mathbb{R}^r$ is a sparse column vector of coefficients identifying the nonzero terms in the dynamic model of the corresponding slow state, $\dot{x}_i = f_i(x)$. Therefore, to determine Ξ, we need to solve the following equation:

$$\dot{X} = \Theta(X)\Xi \tag{16}$$

In brief, Eq. (16) can be solved using a straightforward least-squares routine after neglecting and zeroing all coefficients in Ξ that are smaller than a threshold, λ. The least-squares problem associated with Eq. (16) can most generally be formulated as follows:

$$\Xi = \arg\min_{\Xi'} \left\| \dot{X} - \Theta(X)\Xi' \right\|_2 + \lambda \left\| \Xi' \right\|_1 \tag{17}$$

where Ξ' is a notational substitute for Ξ, and the second term enforces sparsity of Ξ. Practically, we first define the matrix Ξ'' to be the matrix Ξ' with all coefficients with magnitudes below λ set to zero, which is the practical implementation of the L_1 regularization term in Eq. (17). Subsequently, we solve the following least-squares problem

$$\Xi = \arg\min_{\Xi''} \left\| \dot{X} - \Theta(X)\Xi'' \right\|_2 \tag{18}$$

in each iteration using MATLAB's built-in linear solver called with $A\backslash b$ where $A = \dot{X}$ and $b = \Theta(X)$ until the big/nonzero coefficients (larger than λ in each iteration) converge.

To find the Pareto optimal value of the parameter λ that balances model complexity with accuracy, methods such as cross-validation from machine learning may be utilized (Brunton et al., 2016b). In this work, a broad sweep of λ is first used to identify the order of magnitude above which the model has too few nonzero terms to capture the dynamics, resulting in large error values. Subsequently, a narrower sweep in the relevant order of magnitude is used to refine this value of λ. Further refinement was not found to be necessary as a wide range of values from the first refinement yielded identical and optimal models. Once Ξ is found, the overall model is written as the continuous-time differential equation,

$$\dot{x} = \Xi^\top (\Theta(x^\top))^\top$$

where $\Theta(x^\top)$ is not a data matrix but a column vector of symbolic functions of elements of x corresponding to the functions considered in the function library. It is noted that the sampling period of the discrete-time data used to carry out sparse identification affects the accuracy of the finite-difference estimate of the time-derivative, which significantly affects the accuracy of the continuous-time model identified by sparse identification.

3.2. Identifying a model for the slow subsystem

In this work, sparse identification (SI) is used to reconstruct the slow dynamic model for the slow states. The following slow dynamic model is developed to approximate the slow subsystem of Eq. (5).

$$\dot{\hat{x}} = F_{si}(\hat{x}, u) := \hat{f}(\hat{x}) + \hat{g}(\hat{x})u, \qquad \hat{x}(t_0) = x_0 \tag{19}$$

where $\hat{f}(\cdot)$ and $\hat{g}(\cdot)$ are sufficiently smooth vector and matrix functions of dimensions $n \times 1$ and $n \times q$ that approximately capture the functions $f(\cdot)$ and $g(\cdot)$, respectively. Specifically, we first construct the following slow dynamic model using the data set generated with various initial states and $u = 0$ (*i.e.*, steady-state input value):

$$\dot{\hat{x}} = \hat{f}(\hat{x})$$

Subsequently, we use the data set generated with various inputs $u \in U$ and initial states to approximate the function $g(x)$ associated with the input u in the right-hand side of Eq. (6) as $\hat{g}(\hat{x}) = \frac{\dot{\hat{x}} - \hat{f}(\hat{x})}{u}$ where $u \neq 0$. Finally, the model performance is evaluated using unseen data in the testing data set, and is shown to achieve a good representation of $\hat{f}(\cdot)$ and $\hat{g}(\cdot)$ with a desired accuracy.

Remark 3. An alternative method to find $\hat{g}(\cdot)$ is to use a more advanced sparse identification algorithm that expands the function library $\Theta(X)$ to include input terms (*i.e.*, $\Theta(X, u)$) (Brunton et al., 2016a). However, due to the presence of feedback control in this work, this method would require further development before implementation, and the method described above was considered sufficiently accurate.

4. Lyapunov-based MPC using SI models

This section proposes the design of a Lyapunov-based MPC (LMPC) that incorporates the SI model to predict future slow states, followed by a stability analysis of the closed-loop system of Eq. (6) in the succeeding section. Specifically, the stability of the nonlinear system of Eq. (6) under a Lyapunov-based controller associated with the SI model of Eq. (19) is first analyzed. Subsequently, the SI model of Eq. (19) is incorporated into the design of the LMPC under sample-and-hold implementation of the control action to drive the state of the closed-loop system to a small neighborhood around the origin.

4.1. Lyapunov-based control using SI models

For the slow dynamics, it is assumed that there exists a stabilizing control law $u = \Phi_{si}(x) \in U$ such that the origin of the SI slow subsystem of Eq. (19) is rendered asymptotically stable in an open neighborhood $\hat{\phi}_u$ around the origin in the sense that there exist a $\mathcal{C}^1$ control Lyapunov function $\hat{V}(x)$ and four functions, $\hat{\alpha}_1, \hat{\alpha}_2, \hat{\alpha}_3, \hat{\alpha}_4$ of class $\mathcal{K}$ such that $\forall x \in \mathbb{R}^n$:

$$\hat{\alpha}_1(|x|) \leq \hat{V}(x) \leq \hat{\alpha}_2(|x|), \tag{20a}$$

$$\dot{\hat{V}}(x) = \frac{\partial \hat{V}(x)}{\partial x} F_{si}(x, \Phi_{si}(x)) \leq -\hat{\alpha}_3(|x|), \tag{20b}$$

$$\left| \frac{\partial \hat{V}(x)}{\partial x} \right| \leq \hat{\alpha}_4(|x|) \tag{20c}$$

We first characterize the region $\hat{\phi}_u \subset \mathbb{R}^n$ in which the conditions of Eq. (20) are satisfied under the controller $u = \Phi_{si}(x) \in U$. Therefore, starting from inside $\hat{\phi}_u$, the sparse identified slow subsystem of Eq. (19) can be rendered asymptotically stable under the controller $u = \Phi_{si}(x) \in U$. The closed-loop stability region of the sparse identified slow subsystem of Eq. (19) is defined as a level set of the Lyapunov function inside $\hat{\phi}_u$: $\Omega_{\hat{\rho}} := \{x \in \hat{\phi}_u \mid \hat{V}(x) \leq \hat{\rho}\}, \hat{\rho} > 0$. The assumptions of Eq. (8) and Eq. (20) are essentially the stabilizability requirements of the first-principles slow model of Eq. (6) and the sparse identified slow model of Eq. (19), respectively.

Since the dataset for developing the sparse identified slow model of Eq. (6) is obtained from open-loop simulations for $x \in \Omega_\rho$ and $u \in U$, we have $\Omega_{\hat{\rho}} \subseteq \Omega_\rho$. Additionally, there exist positive constants M_{si} and L_{si} such that the following inequalities hold for all $x, x' \in \Omega_{\hat{\rho}}$ and $u \in U$:

$$|F_{si}(x, u)| \leq M_{si} \tag{21a}$$

$$\left| \frac{\partial \hat{V}(x)}{\partial x} F_{si}(x, u) - \frac{\partial \hat{V}(x')}{\partial x} F_{si}(x', u) \right| \leq L_{si}|x - x'| \tag{21b}$$

The following proposition demonstrates that the feedback controller $u = \Phi_{si}(x) \in U$ is able to stabilize the nominal slow subsystem of Eq. (6) in the presence of model mismatch between the nominal subsystem of Eq. (6) and the sparse identified slow model of Eq. (19), provided that the modeling error is sufficiently small.

Proposition 1 (c.f. proposition 2 in Wu et al. (2019a)). *Under the assumption that the closed-loop sparse identified slow subsystem of Eq. (19) is rendered asymptotically stable under the controller $u = \Phi_{si}(x) \in U \; \forall x \in \Omega_{\hat{\rho}}$, if there exists a positive real number $\nu_m < \hat{a}_3(|x|)/\hat{a}_4(|x|)$ that constrains the modeling error $|\nu| = |F(x,u) - F_{si}(x,u)| \leq \nu_m$, $\forall u \in U$ and $\forall x \in \Omega_{\hat{\rho}}$, then the nominal closed-loop system of Eq. (6) under $u = \Phi_{si}(x) \in U$ is also asymptotically stable $\forall x \in \Omega_{\hat{\rho}}$.*

Proof. To prove that the nominal slow subsystem of Eq. (6) can be rendered asymptotically stable $\forall x \in \Omega_{\hat{\rho}}$ under the controller based on the sparse identified model of Eq. (19), we show that $\dot{\hat{V}}$ based on the state of the nominal slow subsystem of Eq. (6) can still be rendered negative under $u = \Phi_{si}(x) \in U$, $\forall\, x \in \Omega_{\hat{\rho}}$.

Based on Eq. (20b) and Eq. (20c), the time-derivative of $\hat{V}$ is computed as follows:

$$\begin{aligned}\dot{\hat{V}} &= \tfrac{\partial \hat{V}(x)}{\partial x} F(x, \Phi_{si}(x)) \\ &= \tfrac{\partial \hat{V}(x)}{\partial x}\big(F_{si}(x, \Phi_{si}(x)) + F(x, \Phi_{si}(x)) - F_{si}(x, \Phi_{si}(x))\big) \\ &\leq -\hat{a}_3(|x|) + \hat{a}_4(|x|)\big(F(x, \Phi_{si}(x)) - F_{si}(x, \Phi_{si}(x))\big) \\ &\leq -\hat{a}_3(|x|) + \hat{a}_4(|x|)\nu_m\end{aligned} \tag{22}$$

If ν_m is chosen to satisfy $\nu_m < \hat{a}_3(|x|)/\hat{a}_4(|x|)$, then it holds that $\dot{\hat{V}} \leq -\tilde{a}_3(|x|) \leq 0$ where $\tilde{a}_3(|x|) = -\hat{a}_3(|x|) + \hat{a}_4(|x|)\nu_m > 0$. This is possible because $\hat{a}_3$ and $\hat{a}_4$ are known functions. For example, if we choose $\hat{a}_3 = a_3|x|$ and $\hat{a}_4 = a_4|x|$, then $\nu_m = a_3/a_4$. As a result, the closed-loop state of the nominal slow subsystem of Eq. (6) converges to the origin under $u = \Phi_{si}(x) \in U$ for all $x_0 \in \Omega_{\hat{\rho}}$. □

After incorporating the SI model of Eq. (19) in the LMPC design, the control actions of the LMPC will be implemented in a sample-and-hold fashion. Hence, the next two propositions demonstrate the sample-and-hold properties of the Lyapunov-based controller $u = \Phi_{si}(x)$. In particular, the following proposition derives an upper bound for the error between the slow states calculated by the nominal slow subsystem of Eq. (6) and the slow states predicted by the SI model of Eq. (19).

Proposition 2 (c.f. proposition 3 in Wu et al. (2019a)). *For the nonlinear system $\dot{x} = F(x,u)$ of Eq. (6) and the SI model $\dot{\hat{x}} = F_{si}(\hat{x},u)$ of Eq. (19) with the same initial condition $x_0 = \hat{x}_0 \in \Omega_{\hat{\rho}}$, there exist a class $\mathcal{K}$ function $f_w(\cdot)$ and a positive constant κ such that the following inequalities hold $\forall x, \hat{x} \in \Omega_{\hat{\rho}}$:*

$$|x(t) - \hat{x}(t)| \leq f_w(t) := \frac{\nu_m}{L}(e^{Lt} - 1) \tag{23a}$$

$$\hat{V}(x) \leq \hat{V}(\hat{x}) + \hat{a}_4\big(\hat{a}_1^{-1}(\hat{\rho})\big)|x - \hat{x}| + \kappa|x - \hat{x}|^2 \tag{23b}$$

Proof. Denoting the error vector between the solutions of the system $\dot{x} = F(x,u)$ and the SI model $\dot{\hat{x}} = F_{si}(\hat{x},u)$ by $e(t) = x(t) - \hat{x}(t)$, the time-derivative of $e(t)$ is obtained as follows:

$$\begin{aligned}|\dot{e}(t)| &= |F(x,u) - F_{si}(\hat{x},u)| \\ &\leq |F(x,u) - F(\hat{x},u)| + |F(\hat{x},u) - F_{si}(\hat{x},u)|\end{aligned} \tag{24}$$

From Eq. (10b), $\forall x, \hat{x} \in \Omega_{\hat{\rho}}$, it is derived that

$$|F(x,u) - F(\hat{x},u)| \leq L|x(t) - \hat{x}(t)| \tag{25}$$

Since the second term $|F(\hat{x},u) - F_{si}(\hat{x},u)|$ in Eq. (24) represents the modeling error, it is bounded by $|\nu| \leq \nu_m$ for all $\hat{x} \in \Omega_{\hat{\rho}}$. Hence, based on Eq. (25) and the bound on the modeling error, $\dot{e}(t)$ is bounded as follows:

$$\begin{aligned}|\dot{e}(t)| &\leq L|x(t) - \hat{x}(t)| + \nu_m \\ &\leq L|e(t)| + \nu_m\end{aligned} \tag{26}$$

Based on the zero initial condition (*i.e.*, $e(0) = 0$), the norm of the error vector can be bounded as follows $\forall x(t), \hat{x}(t) \in \Omega_{\hat{\rho}}$:

$$|e(t)| = |x(t) - \hat{x}(t)| \leq \frac{\nu_m}{L}(e^{Lt} - 1) \tag{27}$$

Subsequently, to derive Eq. (23b) for all $x, \hat{x} \in \Omega_{\hat{\rho}}$, we derive the Taylor series expansion of $\hat{V}(x)$ around $\hat{x}$ as follows:

$$\hat{V}(x) \leq \hat{V}(\hat{x}) + \frac{\partial \hat{V}(\hat{x})}{\partial x}|x - \hat{x}| + \kappa|x - \hat{x}|^2 \tag{28}$$

where κ is a positive real number. Using Eq. (20a) and Eq. (20c), it follows that

$$\hat{V}(x) \leq \hat{V}(\hat{x}) + \hat{a}_4\big(\hat{a}_1^{-1}(\hat{\rho})\big)|x - \hat{x}| + \kappa|x - \hat{x}|^2 \tag{29}$$

This completes the proof of Proposition 2. □

The final proposition below proves that the closed-loop state of the nominal slow subsystem of Eq. (6) remains bounded in $\Omega_{\hat{\rho}}$ for all times, and can be ultimately bounded in a small subset $\Omega_{\rho_{\min}}$ containing the origin under the sample-and-hold implementation of the Lyapunov-based controller $u = \Phi_{si}(x) \in U$.

Proposition 3. *Consider the nominal slow subsystem of Eq. (6) under the controller $u = \Phi_{si}(\hat{x}) \in U$ that is designed to stabilize the SI system of Eq. (19) and meets the conditions of Eq. (20). The controller is implemented in a sample-and-hold fashion, i.e., $u(t) = \Phi_{si}(\hat{x}(t_k))$, $\forall\, t \in [t_k, t_{k+1})$, where $t_{k+1} := t_k + \Delta$. Let ϵ_s, $\epsilon_w > 0$, $\Delta > 0$ and $\hat{\rho} > \rho_{\min} > \rho_{si} > \rho_s$ satisfy*

$$-\hat{a}_3\big(\hat{a}_2^{-1}(\rho_s)\big) + L_{si}M_{si}\Delta \leq -\epsilon_s \tag{30a}$$

$$-\tilde{a}_3\big(\hat{a}_2^{-1}(\rho_s)\big) + L'M\Delta \leq -\epsilon_w \tag{30b}$$

and

$$\rho_{si} := \max\{\hat{V}(\hat{x}(t+\Delta)) \mid \hat{x}(t) \in \Omega_{\rho_s}, u \in U\} \tag{31a}$$

$$\rho_{\min} \geq \rho_{si} + \hat{a}_4\big(\hat{a}_1^{-1}(\hat{\rho})\big) f_w(\Delta) + \kappa(f_w(\Delta))^2 \tag{31b}$$

Then, for any $x(t_k) \in \Omega_{\hat{\rho}} \backslash \Omega_{\rho_s}$, there exists a class $\mathcal{KL}$ function β_x and a class $\mathcal{K}$ function $\bar{\gamma}$ such that the following inequality holds:

$$|x(t)| \leq \beta_x(|x(0)|, t) + \bar{\gamma}(\rho_{\min}) \tag{32}$$

and the slow states $x(t)$ of the nominal subsystem of Eq. (6) is bounded in $\Omega_{\hat{\rho}}$ for all times and ultimately bounded in $\Omega_{\rho_{\min}}$.

Proof. *Part 1*: Assuming $x(t_k) = \hat{x}(t_k) \in \Omega_{\hat{\rho}} \backslash \Omega_{\rho_s}$, we first show that $\hat{V}(\hat{x})$ is decreasing under the controller $u(t) = \Phi_{si}(x(t_k)) \in U$ for $t \in [t_k, t_{k+1})$, where $x(t)$ and $\hat{x}(t)$ denote the solutions of the nominal slow subsystem of Eq. (6) and the SI subsystem of Eq. (19), respectively. The time-derivative of $\hat{V}(\hat{x})$ along the trajectory $\hat{x}(t)$ of the SI model of Eq. (19) for $t \in [t_k, t_{k+1})$ is computed as follows:

$$\begin{aligned}\dot{\hat{V}}(\hat{x}(t)) &= \tfrac{\partial \hat{V}(\hat{x}(t))}{\partial \hat{x}} F_{si}(\hat{x}(t), \Phi_{si}(\hat{x}(t_k))) \\ &= \tfrac{\partial \hat{V}(\hat{x}(t_k))}{\partial \hat{x}} F_{si}(\hat{x}(t_k), \Phi_{si}(\hat{x}(t_k))) \\ &\quad + \tfrac{\partial \hat{V}(\hat{x}(t))}{\partial \hat{x}} F_{si}(\hat{x}(t), \Phi_{si}(\hat{x}(t_k))) \\ &\quad - \tfrac{\partial \hat{V}(\hat{x}(t_k))}{\partial \hat{x}} F_{si}(\hat{x}(t_k), \Phi_{si}(\hat{x}(t_k)))\end{aligned} \tag{33}$$

Combining Eq. (20a) and Eq. (20b), the following inequality is derived:

$$\begin{aligned}\dot{\hat{V}}(\hat{x}(t)) \leq\ & -\hat{a}_3\big(\hat{a}_2^{-1}(\rho_s)\big) + \tfrac{\partial \hat{V}(\hat{x}(t))}{\partial \hat{x}} F_{si}(\hat{x}(t), \Phi_{si}(\hat{x}(t_k))) \\ & - \tfrac{\partial \hat{V}(\hat{x}(t_k))}{\partial \hat{x}} F_{si}(\hat{x}(t_k), \Phi_{si}(\hat{x}(t_k)))\end{aligned} \tag{34}$$

Using the Lipschitz condition of Eq. (21) and the fact that $\hat{x} \in \Omega_{\hat{\rho}}$, $u \in U$, the upper bound of $\dot{\hat{V}}(\hat{x}(t))$ is obtained $\forall\, t \in [t_k, t_{k+1})$:

$$\begin{aligned}\dot{\hat{V}}(\hat{x}(t)) &\leq -\hat{a}_3\left(\hat{a}_2^{-1}(\rho_s)\right) + L_{si}|\hat{x}(t) - \hat{x}(t_k)| \\ &\leq -\hat{a}_3\left(\hat{a}_2^{-1}(\rho_s)\right) + L_{si}M_{si}\Delta\end{aligned} \tag{35}$$

Hence, if Eq. (30a) is satisfied, the following inequality holds $\forall\, \hat{x}(t_k) \in \Omega_{\hat{\rho}} \backslash \Omega_{\rho_s}$, $\forall\, t \in [t_k, t_{k+1})$:

$$\dot{\hat{V}}(\hat{x}(t)) \leq -\epsilon_s \tag{36}$$

Integrating the above differential equation with respect to time over $t \in [t_k, t_{k+1})$, it is derived that $V(\hat{x}(t_{k+1})) \leq V(\hat{x}(t_k)) - \epsilon_s\Delta$. We have shown thus far that for all $\hat{x}(t_k) \in \Omega_{\hat{\rho}} \backslash \Omega_{\rho_s}$, the closed-loop state of the SI slow subsystem of Eq. (19) is bounded inside the closed-loop stability region $\Omega_{\hat{\rho}}$ for all times and moves towards the origin under the controller $u = \Phi_{si}(\hat{x}) \in U$ when implemented in a sample-and-hold fashion.

We note, however, that Eq. (36) may fail to hold when $x(t_k) = \hat{x}(t_k) \in \Omega_{\rho_s}$, which would imply that the state may exit Ω_{ρ_s} within one sampling period. Hence, we design another region $\Omega_{\rho_{si}}$ according to Eq. (31a), which ensures that the closed-loop state $\hat{x}(t)$ of the SI model does not leave $\Omega_{\rho_{si}}$ for all $t \in [t_k, t_{k+1})$, $u \in U$ and $\hat{x}(t_k) \in \Omega_{\rho_s}$ within one sampling period. If the state $\hat{x}(t_{k+1})$ leaves Ω_{ρ_s}, Eq. (36) is satisfied again at $t = t_{k+1}$, reactivating the controller $u = \Phi_{si}(x(t_{k+1}))$ and driving the state towards Ω_{ρ_s} over the next sampling period. As a result, it is shown that the state converges to $\Omega_{\rho_{si}}$ for the closed-loop SI subsystem of Eq. (19) for all $\hat{x}_0 \in \Omega_{\hat{\rho}}$. In Part 2, we show that the closed-loop state of the nominal slow subsystem of Eq. (6) can also be bounded in $\Omega_{\hat{\rho}}$ for all times and ultimately bounded in a small neighborhood around the origin under the sample-and-hold implementation of the controller $u = \Phi_{si}(x) \in U$.

Part 2: Repeating the analysis performed for the SI subsystem of Eq. (19), we first assume $x(t_k) = \hat{x}(t_k) \in \Omega_{\hat{\rho}} \backslash \Omega_{\rho_s}$ and compute the time-derivative of $\hat{V}(x)$ for the nominal slow subsystem of Eq. (6) (*i.e.*, $\dot{x} = F(x, u)$) as follows:

$$\begin{aligned}\dot{\hat{V}}(x(t)) &= \tfrac{\partial \hat{V}(x(t))}{\partial x} F(x(t), \Phi_{si}(x(t_k))) \\ &= \tfrac{\partial \hat{V}(x(t_k))}{\partial x} F(x(t_k), \Phi_{si}(x(t_k))) \\ &\quad + \tfrac{\partial \hat{V}(x(t))}{\partial x} F(x(t), \Phi_{si}(x(t_k))) \\ &\quad - \tfrac{\partial \hat{V}(x(t_k))}{\partial x} F(x(t_k), \Phi_{si}(x(t_k)))\end{aligned} \tag{37}$$

From Eq. (22), $\frac{\partial \hat{V}(x(t_k))}{\partial x} F(x(t_k), \Phi_{si}(x(t_k))) \leq -\tilde{a}_3\left(x(t_k)\right)$ holds for all $x \in \Omega_{\hat{\rho}} \backslash \Omega_{\rho_s}$ where $\tilde{a}_3(\cdot)$ was defined at the end of Proposition 1. Using Eq. (8a) and the Lipschitz definition in Eq. (10), the following inequality is obtained for $\dot{\hat{V}}(x(t))$, $\forall\, t \in [t_k, t_{k+1})$ and $\forall\, x(t_k) \in \Omega_{\hat{\rho}} \backslash \Omega_{\rho_s}$:

$$\begin{aligned}\dot{\hat{V}}(x(t)) &\leq -\tilde{a}_3\left(\hat{a}_2^{-1}(\rho_s)\right) + \tfrac{\partial \hat{V}(x(t))}{\partial x} F(x(t), \Phi_{si}(x(t_k))) \\ &\quad - \tfrac{\partial \hat{V}(x(t_k))}{\partial x} F(x(t_k), \Phi_{si}(x(t_k))) \\ &\leq -\tilde{a}_3\left(\hat{a}_2^{-1}(\rho_s)\right) + L'|x(t) - x(t_k)| \\ &\leq -\tilde{a}_3\left(\hat{a}_2^{-1}(\rho_s)\right) + L'M\Delta\end{aligned} \tag{38}$$

Consequently, if Eq. (30b) is satisfied, the following inequality holds $\forall\, x(t_k) \in \Omega_{\hat{\rho}} \backslash \Omega_{\rho_s}$, $\forall\, t \in [t_k, t_{k+1})$:

$$\dot{\hat{V}}(x(t)) \leq \quad -\epsilon_w \tag{39}$$

Integrating the above differential equation with respect to time between any two points in $[t_k, t_{k+1})$, it is derived for all $x(t_k) \in \Omega_{\hat{\rho}} \backslash \Omega_{\rho_s}$:

$$\hat{V}(x(t_{k+1})) \leq V(x(t_k)) - \epsilon_w\Delta \tag{40}$$

$$\hat{V}(x(t)) \leq \hat{V}(x(t_k)), \forall\, t \in [t_k, t_{k+1}) \tag{41}$$

Therefore, the state of the closed-loop system of Eq. (6) remains in $\Omega_{\hat{\rho}}$ for all times. Furthermore, it follows that the controller $u = \Phi_{si}(x)$ is still able to drive the state of the nominal slow subsystem of Eq. (6) towards the origin within every sampling period. Moreover, if $x(t_k) \in \Omega_{\rho_s}$, it was already shown in Part 1 that the state of the SI model of Eq. (19) is maintained in $\Omega_{\rho_{si}}$ for one sampling period. Considering the bounded modeling error between the state of the SI model of Eq. (19) and the state of the nominal slow subsystem of Eq. (6) given by Eq. (30a), there exists a compact set $\Omega_{\rho_{\min}} \supset \Omega_{\rho_{si}}$ that satisfies Eq. (31b) such that the state of the nominal slow subsystem of Eq. (6) remains within $\Omega_{\rho_{\min}}$ during one sampling period if the state of the SI model of Eq. (19) is bounded in $\Omega_{\rho_{si}}$. If the state $x(t)$ enters $\Omega_{\rho_{\min}} \backslash \Omega_{\rho_s}$, we have shown that Eq. (41) holds, and thus, the state will be driven towards the origin again under $u = \Phi_{si}(x)$ during the next sampling period, ultimately bounding the closed-loop slow subsystem in $\Omega_{\rho_{\min}}$. Therefore, based on the continuity of the Lyapunov function $\hat{V}$, there exist a class $\mathcal{KL}$ function β_x and a class $\mathcal{K}$ function $\bar{\gamma}$ such that if $x_0 \in \Omega_{\hat{\rho}}$, then $x(t) \in \Omega_{\hat{\rho}}$, $\forall\, t \geq t_0$ and

$$|x(t)| \leq \beta_x(|x(0)|, t) + \bar{\gamma}(\rho_{\min}) \tag{42}$$

□

4.2. Lyapunov-based MPC (LMPC) formulation

The LMPC is based on the Lyapunov-based controller $\Phi_{si}(x)$. The controller $\Phi_{si}(x)$ is used to define a stability constraint for the LMPC controller. This ensures that the LMPC controller inherits the stability and robustness properties of the Lyapunov-based controller $\Phi_{si}(x)$. The LMPC controller is given by the following optimization problem:

$$\mathcal{J} = \min_{u \in S(\Delta)} \int_{t_k}^{t_{k+N}} L(\tilde{x}(t), u(t))\, \mathrm{d}t \tag{43a}$$

$$\text{s.t.} \quad \dot{\tilde{x}}(t) = F_{si}(\tilde{x}(t), u(t)) \tag{43b}$$

$$u(t) \in U, \ \forall\, t \in [t_k, t_{k+N}) \tag{43c}$$

$$\tilde{x}(t_k) = x(t_k) \tag{43d}$$

$$\dot{\hat{V}}(x(t_k), u) \leq \dot{\hat{V}}(x(t_k), \Phi_{si}(x(t_k))), \ \text{if } x(t_k) \in \Omega_{\hat{\rho}} \backslash \Omega_{\rho_{si}} \tag{43e}$$

$$\hat{V}(\tilde{x}(t)) \leq \rho_{si}, \ \forall\, t \in [t_k, t_{k+N}), \ \text{if } x(t_k) \in \Omega_{\rho_{si}} \tag{43f}$$

where $\tilde{x}$ denotes the predicted trajectory of the slow states, $S(\Delta)$ is the set of piece-wise constant functions with period Δ, and N is the number of sampling periods within the prediction horizon. $\dot{\hat{V}}(x, u)$ denotes $\frac{\partial \hat{V}(x)}{\partial x}(F_{si}(x, u))$. Let $u = u^*(t)$, $t \in [t_k, t_{k+N})$ denote the optimal input trajectory calculated by the LMPC over the entire prediction horizon. It is noted that only the first control action of the computed sequence, $u^*(t_k)$, which corresponds to the first sampling period of the prediction horizon, is applied over the first sampling period, and the LMPC is resolved at the next sampling time.

In the optimization problem of Eq. (43), the objective function of Eq. (43a) is equal to the integral of $L(\tilde{x}(t), u(t))$ over the entire prediction horizon. The constraint of Eq. (43b) is the approximate slow model of Eq. (19) that is used to predict the slow states of the closed-loop slow subsystem. Eq. (43c) specifies the input constraints to be applied over the entire prediction horizon. Eq. (43d) defines the initial condition $\tilde{x}(t_k)$ of Eq. (43b), which is

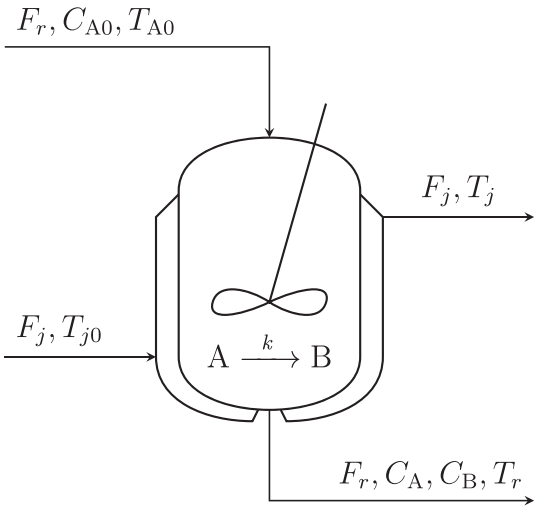

Fig. 1. A continuous-stirred tank reactor with jacket.

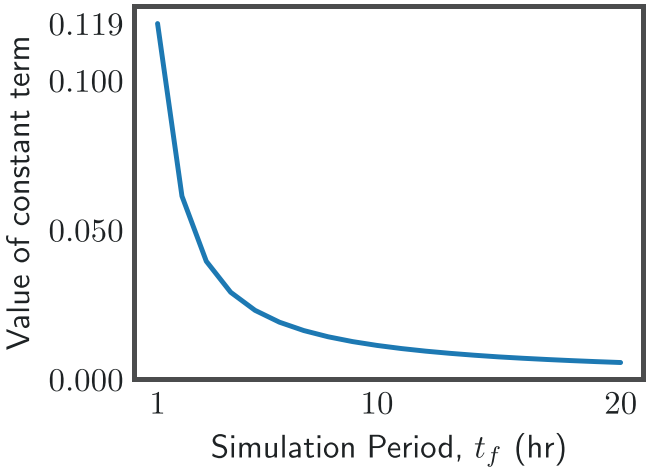

Fig. 2. Plot of the constant residual term in Eq. (58) as a function of the total simulation duration t_f used for data generation.

the slow state measurement at $t = t_k$. The first Lyapunov constraint of Eq. (43e) guarantees that the closed-loop state moves towards the origin if $x(t_k) \in \Omega_{\hat{\rho}} \backslash \Omega_{\rho_{si}}$. However, if $x(t_k)$ enters $\Omega_{\rho_{si}}$, the states predicted by the approximate slow model of Eq. (43b) will be maintained in $\Omega_{\rho_{si}}$ over the entire prediction horizon.

5. Closed-loop stability analysis

The closed-loop stability of the singularly perturbed system of Eq. (1) under the LMPC of Eq. (43) is established in the following theorem under appropriate conditions.

Theorem 1. *Consider the system of Eq. (1) in closed-loop with u^* computed by the LMPC of Eq. (43) based on the Lyapunov-based controller $\Phi_{si}(x)$ that satisfies the conditions of Eq. (20). Let Assumptions 1 and 2 and the conditions of Propositions 1–3 hold. Then there exist functions β_x, β_y of class $\mathcal{KL}$, a pair of positive real numbers (δ, d) and $\exists\, \epsilon^* > 0$ such that if $\max\{|x(0)|, |y(0)|\} \leq \delta$ and $\epsilon \in (0, \epsilon^*]$, then $\forall\, t \geq 0$,*

$$|x(t)| \leq \beta_x(|x(0)|, t) + \bar{\gamma}(\rho_{\min}) + d \tag{44}$$

$$|y(t)| \leq \beta_y\left(|y(0)|, \frac{t}{\epsilon}\right) + d \tag{45}$$

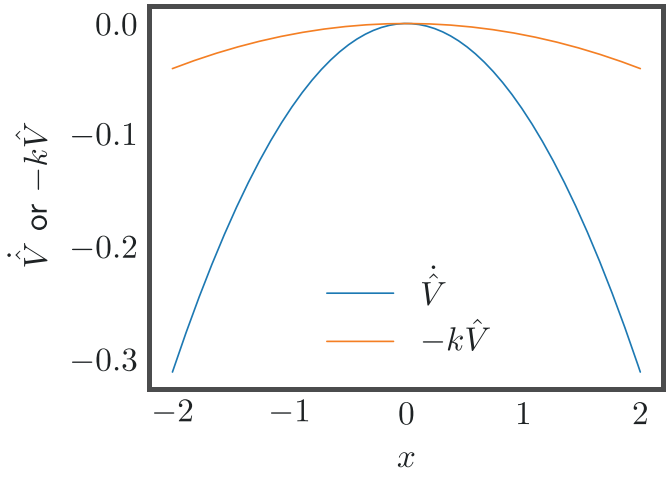

Fig. 3. Plot of $\dot{\hat{V}}$ and $-k\hat{V}$.

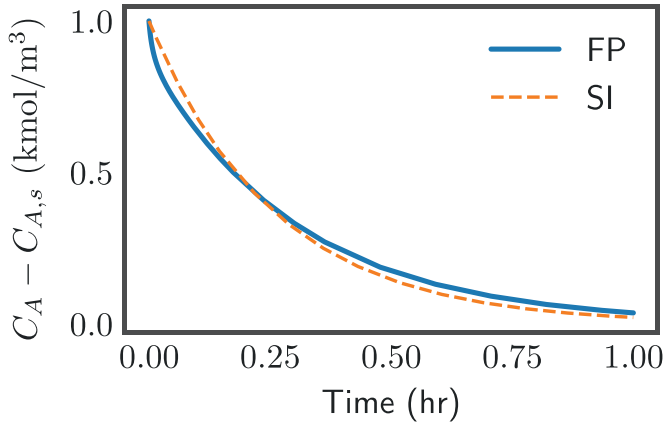

Fig. 4. Comparison of the slow state as computed by the first-principles slow subsystem (solid line) of Eq. (52) and predicted by the SI slow subsystem (dashed line) of Eq. (59) with $u = 0$.

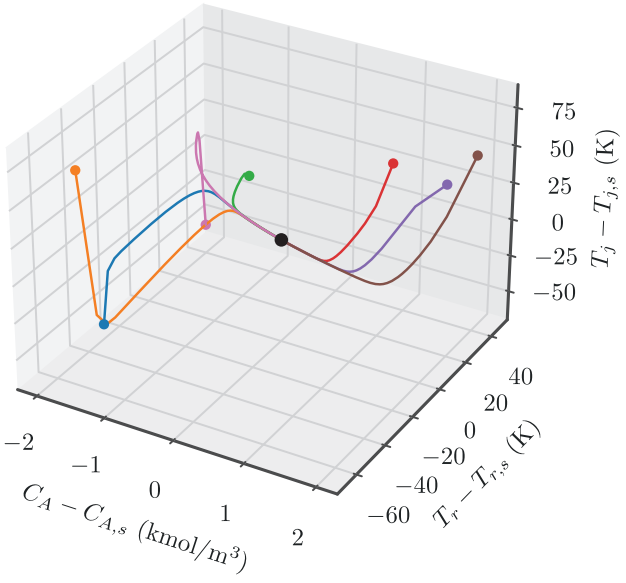

Fig. 5. State-space trajectories for the CSTR in closed-loop under the LMPC utilizing the SI slow model for a range of initial conditions. Each line represents a trajectory from an initial condition marked by a colored dot of the corresponding color. The black dot is the origin.

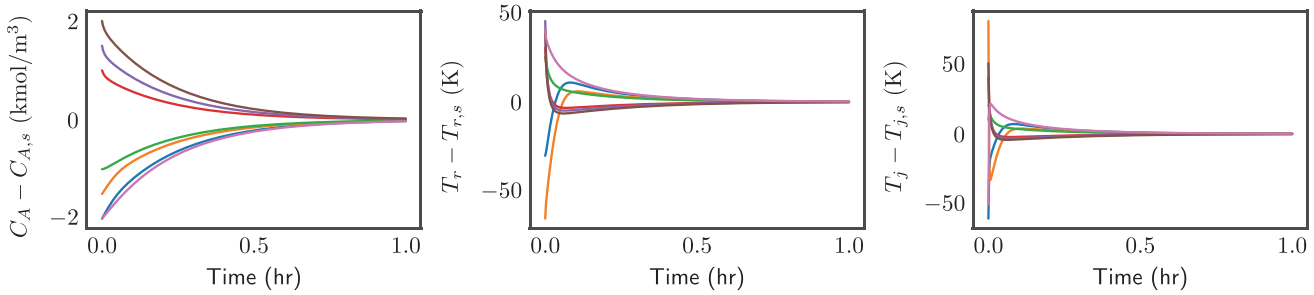

Fig. 6. State profiles for the CSTR in closed-loop under the LMPC utilizing the SI slow model for a range of initial conditions, each line representing a trajectory from a different initial condition.

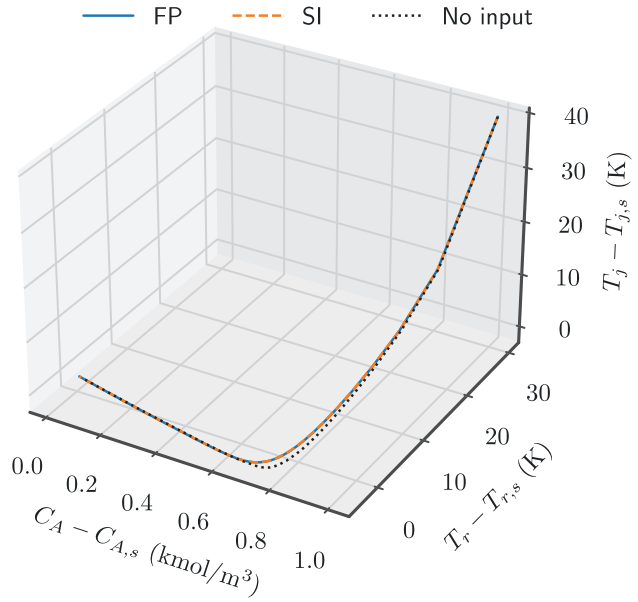

Fig. 7. State-space profiles for the CSTR in closed-loop under the LMPC utilizing the first-principles model (blue line) and the SI slow model (orange line) and in open-loop with $u = 0$ (black line). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Proof. The closed-loop system takes the following form after substituting the optimal control action u^* into Eq. (1).

$$\dot{x} = f_1(x, z, u^*, \epsilon) \tag{46a}$$

$$\epsilon \dot{z} = f_2(x, z, \epsilon) \tag{46b}$$

By setting $\epsilon = 0$, we have

$$\dot{x} = f_1(x, z, u^*, 0) \tag{47a}$$

$$0 = f_2(x, z, 0) \tag{47b}$$

Since Eq. (47b) has a unique, isolated solution $z_s = \hat{f}_2(x)$, following Eq. (4), Eq. (47) can be written in the following form:

$$\dot{x} = f_1(x, \hat{f}_2(x), u^*, 0) \tag{48}$$

When $x(t_k) \in \Omega_{\hat{\rho}} \backslash \Omega_{\rho_{si}}$, the constraint of Eq. (43e) requires that the Lyapunov function value decrease at least at the rate under $u = \Phi_{si}(x)$. As a result, the time-derivative of Lyapunov function $\dot{V}$ under $u = u^*$ is rendered negative. Based on the results in Proposition 3, the state of the slow subsystem of Eq. (46a) will approach the origin and enter $\Omega_{\rho_{si}}$ within finite sampling steps provided that the modeling error is sufficiently small. After $x(t_k)$ enters $\Omega_{\rho_{si}}$, the constraint of Eq. (43f) maintains the predicted state within $\Omega_{\rho_{si}}$ afterwards. Since the modeling error and sampling period are sufficiently small, we have shown in Proposition 3 that the true state of Eq. (46a) can be bounded in $\Omega_{\rho_{\min}}$, which is a slightly larger set containing $\Omega_{\rho_{si}}$. Therefore, for any initial state $x_0 \in \Omega_{\hat{\rho}}$, LMPC ensures that the state $x(t)$ of the closed-loop slow subsystem of Eq. (46a) is bounded in $\Omega_{\hat{\rho}}$ for all times, and satisfies the bound of Eq. (32) in Proposition 3.

Subsequently, by letting $\tau = \frac{t}{\epsilon}$, $y = z - \hat{f}_2(x)$ and $\epsilon = 0$, we obtain the closed-loop fast subsystem:

$$\frac{dy}{d\tau} = f_2(x, \hat{f}_2(x) + y, 0) \tag{49}$$

Note that the origin of the closed-loop system of Eq. (49) is assumed to be globally asymptotically stable such that Eq. (7) holds for any $y(0) \in \mathbb{R}^p$ in Assumption 2. Therefore, the closed-loop system of Eq. (46) satisfies all the assumptions for Theorem 1 in Christofides and Teel (1996), which implies that there exist class $\mathcal{KL}$ functions $\beta_x(\cdot)$ and $\beta_y(\cdot)$, positive real numbers (δ, d), and $\epsilon^* > 0$ such that if $\max\{|x(0)|, |y(0)|\} \leq \delta$ and $\epsilon \in (0, \epsilon^*]$, then, the slow and fast system states are bounded by Eqs. (44) and (45). □

6. Application to a chemical process example

6.1. Process description

We demonstrate the application of the LMPC of Eq. (43) based on the sparse-identified slow subsystem using a chemical process example. We consider a perfectly-mixed, non-isothermal CSTR as shown in Fig. 1.

A single, endothermic, irreversible reaction of the form,

$$\mathrm{A} \xrightarrow{\mathrm{k}} \mathrm{B} \tag{50}$$

occurs in the CSTR. The concentration of reactant A in the reactor is denoted by C_A. Assuming the vessel has a constant holdup, the volume of the liquid in the reactor is represented by V_r. The temperature of the reactor contents is denoted by T_r. The feed to the reactor contains pure species A with molar concentration C_{A0}, at a flow rate F, and temperature T_{A0}. Owing to the endothermic reaction taking place in the reactor, energy must be provided to the reactor via a jacket. The heating jacket has a volume V_j with heat transfer fluid at an inlet temperature of T_{j0} being added to it at a flow rate F_j. The reactor contents and the heat transfer fluid have constant densities of ρ_m and ρ_j, respectively, and have constant heat capacities of $c_{p,m}$ and $c_{p,j}$, respectively. ΔH_r denotes the enthalpy of the reaction, U represents the heat transfer coefficient, and A_r is the heat transfer contact area between the reactor and the jacket. The rate constant k in Eq. (50) is assumed to be of the form,

$$k = k_0 \exp\left(\frac{-E}{RT_r}\right) \tag{51}$$

where k_0, R, and E denote the pre-exponential constant, ideal gas constant, and activation energy of the reaction, respectively. Under

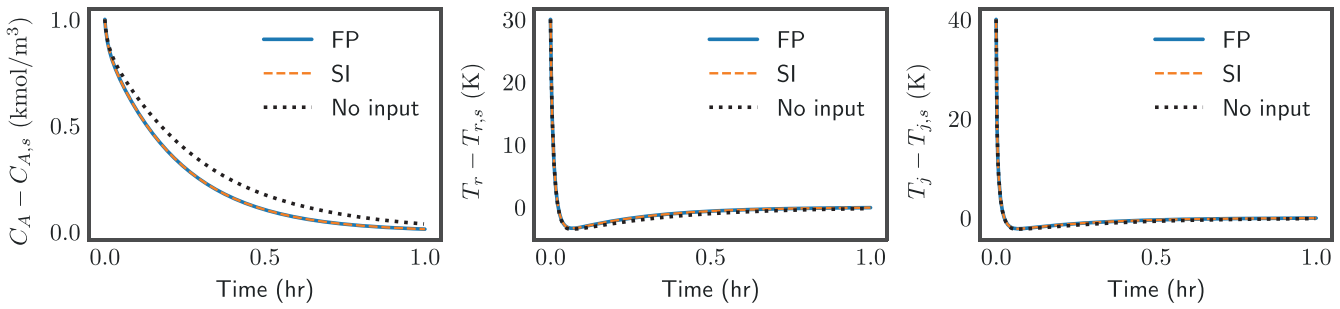

Fig. 8. State profiles for the CSTR in closed-loop under the LMPC utilizing the first-principles model (blue line) and the SI slow model (orange line) and in open-loop with $u = 0$ (black line). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

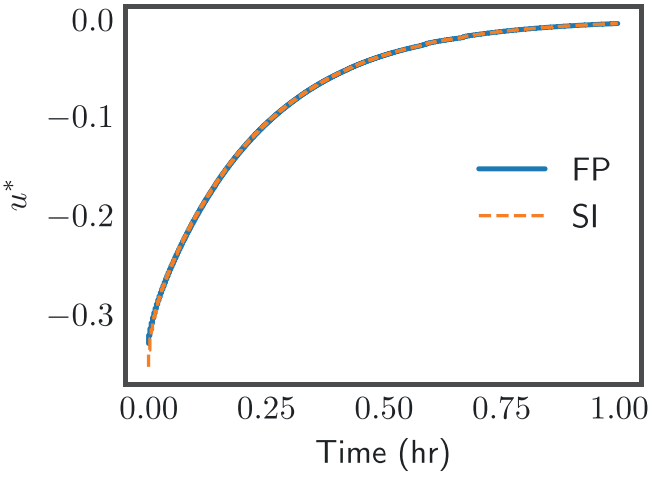

Fig. 9. Input profiles for the CSTR in closed-loop under the LMPC utilizing the first-principles model (blue line) and the SI slow model (orange line). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

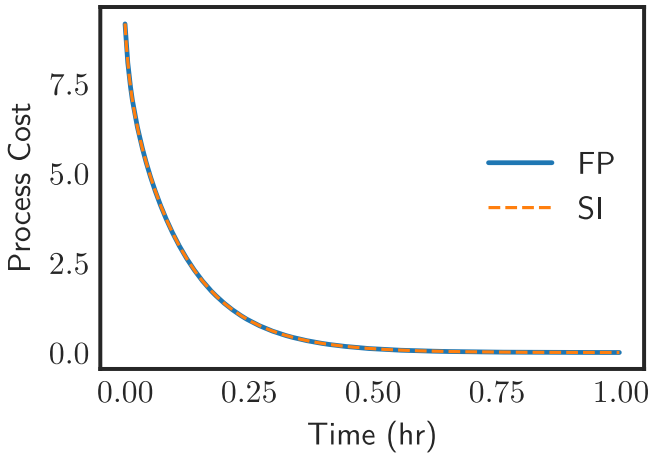

Fig. 10. Cost function of LMPC for the CSTR in closed-loop under the LMPC utilizing the first-principles model (blue line) and the SI slow model (orange line). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 1
Parameter values for chemical process example.

$V_r = 1.0\text{m}^3$	$k_0 = 3.36 \times 10^6\text{h}^{-1}$
$V_j = 0.08\text{m}^3$	$E = 8.0 \times 10^3$ kcal kg^{-1}
$A_r = 6.0\text{m}^3$	$T_{A0} = 310.0\text{K}$
$U = 1000.0\text{kcalh}^{-1}\text{m}^{-2}\text{K}^{-1}$	$T_{j0} = 357.5\text{K}$
$R = 1.987\text{kcal kmol}^{-1}\text{K}^{-1}$	$\rho_m = 900.0\text{kg m}^{-3}$
$\Delta H_r = 5.4 \times 10^4\text{kcal mol}^{-1}$	$\rho_j = 800.0\text{kg m}^{-3}$
$c_{p,m} = 0.231\text{kcal kg}^{-1}\text{K}^{-1}$	$F_r = 3.0\text{m}^3\text{h}^{-1}$
$c_{p,j} = 0.200\text{kcal kg}^{-1}\text{K}^{-1}$	$F_j = 20.0\text{m}^3\text{h}^{-1}$
$C_{A0,s} = 3.75\text{kmol m}^{-3}$	$C_{A,s} = 2.54\text{kmol m}^{-3}$
$T_{r,s} = 274.4\text{K}$	$T_{j,s} = 303.3\text{K}$

these modeling assumptions, the dynamic equations governing the behavior of the reactor are given by the following material and energy balances:

$$V_r \frac{dC_A}{dt} = F_r(C_{A0} - C_A) - k_0 e^{-E/RT_r} C_A V_r \tag{52}$$

$$V_r \frac{dT_r}{dt} = F_r(T_{A0} - T_r) + \frac{(-\Delta H_r)}{\rho_m c_{p,m}} k_0 e^{-E/RT_r} C_A V_r + \frac{UA_r}{\rho_m c_{p,m}}\left(T_j - T_r\right) \tag{53}$$

$$V_j \frac{dT_j}{dt} = F_j(T_{j0} - T_j) - \frac{UA_r}{\rho_j c_{p,j}}\left(T_j - T_r\right) \tag{54}$$

with the process parameters' values given in Table 1.

While we may define $\epsilon = V_r/V_j$ to rewrite Eq. (54) in the standard singularly perturbed form, in which case both temperatures have fast dynamics relative to the concentration dynamics, this can also be deduced from the data generation outlined in the next subsection.

The initial operating point of the CSTR is its steady-state $(C_{A,s}, T_{r,s}, T_{j,s}) = (2.54\text{kmol m}^{-3},\ 274.4\text{K}, 303.3K)$ and $C_{A0,s} = 3.75\text{kmol m}^{-3}$. The manipulated input variable is the feed concentration of reactant A, represented by the deviation variable $\Delta C_{A0} = C_{A0} - C_{A0,s}$ and bounded as per $|\Delta C_{A0}| \leq 3.5\text{kmol m}^{-3}$. Consequently, the states and manipulated input of the closed-loop system in deviation form are represented by

$$x = C_A - C_{A,s} \tag{55}$$

and

$$u = \Delta C_{A0} \tag{56}$$

respectively. Hence, the origin of the state-space, given by $(x_s^*, u_s^*) = (0, 0)$, is the equilibrium point of the system.

It is desired to apply the SI-based LMPC of Eq. (43) to maintain the operation of the CSTR at the equilibrium point $(C_{A,s}, T_{r,s}, T_{j,s})$ by manipulating the feed concentration ΔC_{A0}. For the simulation of the ODE system of Eqs. (52) to (54), the explicit Euler method is used to numerically integrate the equations with an integration time step of $h_c = 0.1$seconds. Since the optimization problem in the LMPC of Eq. (43) is nonlinear, it is solved using PyIpopt, the python interface of the IPOPT software package (Wächter and Biegler, 2006), with a sampling period of $\Delta = 10$seconds.

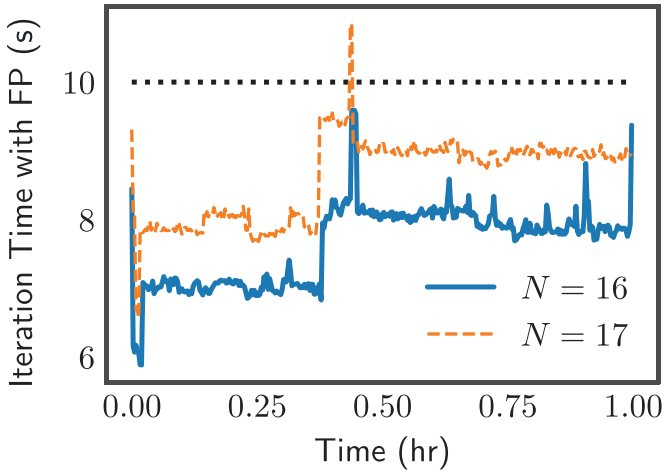

Fig. 11. Iteration time of each MPC step for prediction horizon lengths of $N = 16$ (blue line) and $N = 17$ (orange line) when the LMPC utilizes the first-principles process model, where the black dotted line represents the sampling time $\Delta = 10$ s. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

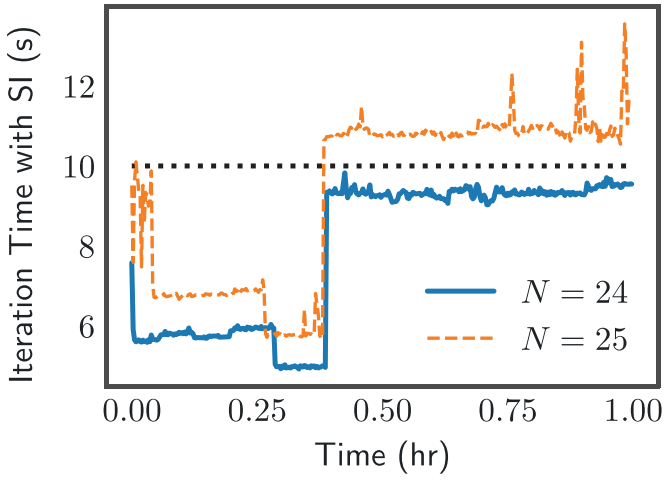

Fig. 12. Iteration time of each MPC step for prediction horizon lengths of $N = 23$ (blue line) and $N = 24$ (orange line) when the LMPC utilizes the sparse-identified slow model, where the black dotted line represents the sampling time $\Delta = 10$ s. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

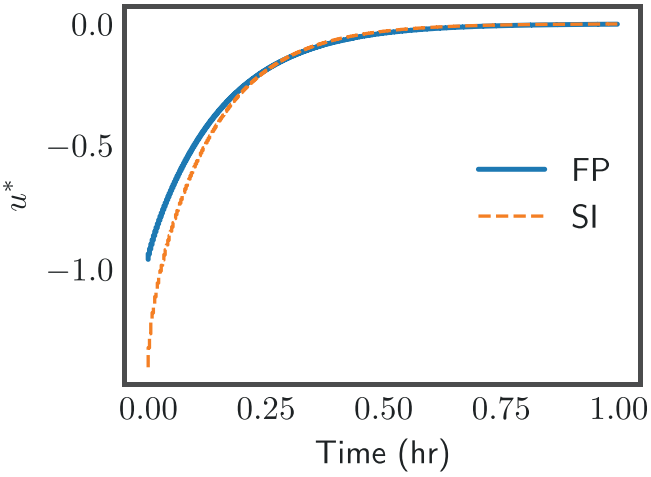

Fig. 13. Input profiles for the CSTR in closed-loop under the LMPC utilizing the first-principles model with $N = 16$ (blue line) and the SI slow model with $N = 24$ (orange line). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

6.2. Data generation and model development

To design the LMPC of Eq. (43), the slow subsystem must be first reconstructed as an ODE to be incorporated as the process model. To identify the slow subsystem using data, a range of 10 different initial conditions are chosen with $C_A(0)$ taking values between 0mol m^{-3} and 9.0mol m^{-3} in intervals of 1.0mol m^{-3}, $T_r(0)$ taking values between 280 K and 370 K in intervals of 10 K, and $T_j(0)$ varying between 300 K and 390 K in intervals of 10 K. The system of differential equations described by Eqs. (52) to (54) is numerically integrated for each set of initial conditions from the initial time of 0.0 hr to 1.0 hr with a step size of 1×10^{-6} hr to simulate the experimental/industrial process. The generated data is then sampled with a sampling period larger than the step size to simulate data collection via sensors and measuring instruments. This sampling is conducted with a sampling period of 0.005 hr since any larger sampling period leads to inaccurate estimates of the time-derivative, causing the sparse identification algorithm to fail for this system. The sampled data from the 10 runs is finally combined into the data matrix, X. A plot of X and/or the gradient of X *versus* time clearly indicates that both temperature variables, T_r and T_j, are the fast states, while C_A is the only slow state in this system. Examples of such plots can be found in Figures. 7, 8 and 6 in the simulation results (Section 6.3). Therefore, we only construct a model for the concentration C_A using sparse identification, as it is the slow subsystem.

The sparse identification algorithm was applied to the concentration data from the open-loop simulations. Ten iterations were completed, and the nonzero coefficients were confirmed to have converged within five iterations. The coefficient threshold λ was fine-tuned to a value of 2.0 to yield the following slow model,

$$\frac{dC_A}{dt} = 9.93101 - 3.85957C_A \tag{57}$$

However, Eq. (57) needed to be converted to deviation variables and the input accounted for. Hence, using Eq. (55), Eq. (57) was written as

$$\frac{dx}{dt} = 9.93101 - 3.85957(x + C_{A,s}) = 0.11901 - 3.85957x \tag{58}$$

In Eq. (58), the constant term, 0.11901, was due to numerical inaccuracies resulting from modeling via sparse identification and the lack of data near the origin of the system in deviation variables (which corresponds to data near the steady-state of the original system). The data generation was carried out over a 1-hour simulation duration, within which the system did not reach very close to its steady-state. However, the simulation length can be increased beyond 1 h to increase the amount of data near the steady-state, which would improve model performance near the origin. The value of this constant residual term in Eq. (58) was found to monotonically decrease with increasing simulation length for the data generation step as shown in Fig. 2. If the data generation is carried out over a simulation length of 20 hours, the value of the residual term reduces to 0.00562783. The optimal sparse identification model built from data generated over an infinitely long simulation would not contain the constant term. Furthermore, we know that the origin is an equilibrium point for the system in deviation variables. Therefore, the constant term in Eq. (58) was neglected. Additionally, the input term must be present in the closed-loop model. The procedure outlined in Section 3.2 was used to calculate $\hat{g}(x)$, which was approximately between 3 and 4. Comparing this result to the first-principles model, where the term corresponding to the manipulated input was known to be $3u$, $\hat{g}(x)$ is taken to be equal

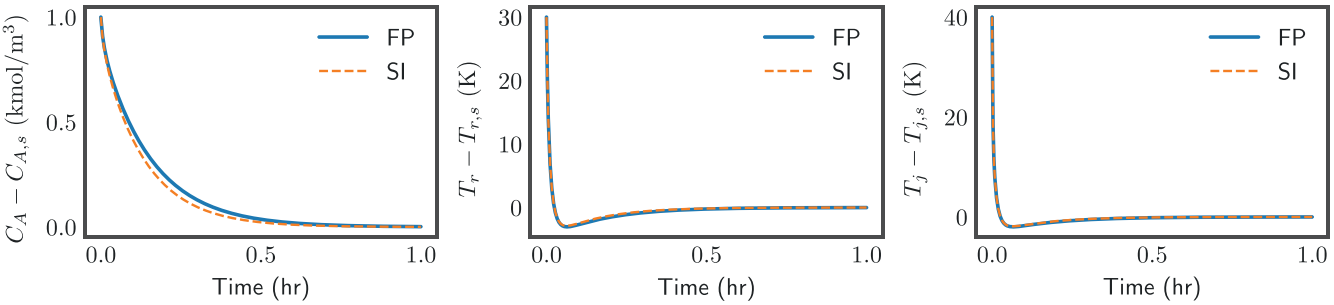

Fig. 14. State profiles for the CSTR in closed-loop under the LMPC utilizing the first-principles model with $N = 16$ (blue line) and the SI slow model with $N = 24$ (orange line). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

to 3. Hence, the sparse-identified slow subsystem was identified as

$$\frac{dx}{dt} = -3.85957x + 3u \tag{59}$$

While the above procedure required partial knowledge of the first-principles model of the process, the stability results of Section 5 merely assumed an input-affine model as given in Eq. (19). Hence, we could alternatively add a manipulated input term of u rather than $3u$ to Eq. (59), and the stability results would hold, while the simulation results would scale accordingly. Furthermore, the coefficient associated with u could also be found more accurately using steady-state data for different input sequences and initial conditions. Henceforth, Eq. (59) will be referred to as the SI slow subsystem or SI slow model for the CSTR system.

After the model identification, the Lyapunov functions for the nominal slow subsystem and the SI slow subsystem are chosen to be identical *i.e.,* $V(x) = \hat{V}(x) = 0.01x^2$. The region $\hat{\phi}_u$ where $\dot{\hat{V}} < -k\hat{V}$ under the controller $u = \Phi_{si}(x) \in U$ is analyzed in Fig. 3. The range $x \in [-2, 2]$ is considered because the steady-state concentration is 2.54kmol m^{-3}, and concentrations may not be negative. As the Lyapunov function in this case is a quadratic function, its level sets are represented by intervals in x. Hence, the closed-loop stability region $\Omega_{\hat{\rho}}$ for the reactor system described by the SI slow subsystem is represented by the largest level set of $\hat{V}$ in $\hat{\phi}_u$. From Fig. 3, it is observed that $\dot{\hat{V}} < -k\hat{V}$ is satisfied for the entire range of x. Therefore, the closed-loop stability region $\Omega_{\hat{\rho}}$ is also characterized as the entire range of $x \in [-2, 2]$. Considering $|x| < 0.05$ to be sufficiently close to the origin, a value of $2.5 \times 10^{-}5$ was calculated for ρ_{si}. The cost function in Eq. (43a) is chosen to be $L(x, u) = |x|^2_{Q_1} + |u|^2_{Q_2}$ where $Q_1 = 10$ and $Q_2 = 1$, such that L achieves its minimum at the origin of the closed-loop system.

6.3. Simulation results

First, simulations are conducted for the closed-loop system under the MPC using the nominal slow subsystem and the SI slow subsystem as the process model of Eq. (43b), separately. While, in practice, if solely data is available, only the sparse identification approach can be used, the nominal slow subsystem is used to represent the experimental process. Hence, the LMPC utilizing the nominal slow subsystem yields the ideal performance that an LMPC with a data-driven process model can achieve and be compared against. Fig. 4 shows that the open-loop state trajectory of the slow state predicted by the sparse-identified concentration model of Eq. (59) with $u = 0$ is in close agreement with the first-principles concentration model of Eq. (52), for a fixed time interval and the same initial condition of $x_0 = 1 \in \Omega_{\hat{\rho}}$. As a result, it can be concluded that the sparse-identified model of Eq. (59) can be considered a satisfactory approximation of the first-principles process model of Eq. (52). Consequently, the sparse-identified slow model of Eq. (59) is used as the process model in the LMPC of Eq. (43).

Figures. 5 and 6 depict the closed-loop states of the CSTR for a range of initial conditions $x_0 \in \Omega_{\hat{\rho}}$ under the LMPC utilizing the SI slow model. It is observed that the states converge to a small neighborhood containing the origin, $\Omega_{\rho_{\min}}$, for all initial conditions studied. Therefore, the LMPC with the SI slow model can be considered adequate to stabilize the CSTR system due to the sufficiently small modeling error.

The closed-loop states and manipulated input profiles of the CSTR system of Eqs. (52) to (54) under the LMPC are shown in Figs. 7, 8 and 9. Figure 7 compares the state trajectories for the closed-loop system from an initial condition of $(C_A - C_{A,s}, T_r - T_{r,s}, T_j - T_{j,s}) = (1\text{kmol m}^{-3}, 30\text{K}, 40\text{K})$ when the different process models are used in the controller. An MPC prediction horizon length of $N = 5$ is chosen for both cases. In both simulations, it is observed that the state trajectory is driven to $\Omega_{\rho_{\min}}$ under the controller faster than in the open-loop scenario without a controller. Analyzing the time-varying states more closely as shown in Fig. 8, it is demonstrated that the evolution of the states is nearly identical for both the nominal slow model and the SI slow model. Fig. 9 shows the manipulated input profiles for the closed-loop system, which are observed to be within the range of permissible u. The results of Figs. 7, 8, 9 were also reproduced for other initial conditions, implying the LMPC incorporating the SI slow model is able to stabilize the CSTR system and drive the states to the origin nearly as efficiently as the LMPC utilizing the nominal slow model. Furthermore, when the integral of the cost function of the LMPC $\int_{t=0}^{t_p} L(x(\tau), u(\tau))\, d\tau$ is calculated over the simulation period $t_p = 1$hr, it is obtained that $L = 330.2$ and $L = 330.0$ for the LMPC utilizing the first-principles model and the SI slow model, respectively. The negligible difference indicates that the closed-loop performance under both models is similar with respect to energy as well as speed of convergence to the origin. The value of the cost $\mathcal{J}$ over the simulation period of 1 hr is shown in Fig. 10, indicating that there is no significant difference in the states and manipulated inputs computed by the LMPC using either process model.

6.4. Effect of process model selection on computational time and maximum allowable prediction horizon length

The choice of the process model for the LMPC has a significant effect on the total as well as per iteration computational time of the MPC, directly limiting the maximum prediction horizon length that can be implemented. Due to the increased complexity of the first-principles process model, the computational time required to solve the LMPC optimization problem for the LMPC incorporating the first-principles slow model is higher for any given prediction horizon length. Consequently, the longest prediction horizon that can be implemented in the controller is lower for the LMPC with

Table 2
Maximum allowable prediction horizon length and corresponding integral of cost function over simulation period t_p.

LMPC Process Model	Max allowable N	Integral of Cost Function
First-principles slow model	16	243.8
Sparse-identified slow model	24	225.9

the first-principles model, leading to a higher total of the integral of the LMPC cost function over the simulation period.

Figs. 11 and 12 depict the iteration time required for each of the 360 LMPC steps overs the entire simulation period of 3600 s with a sampling period Δ of 10 s for the LMPC under the first-principles model and the SI slow model, respectively, while using the initial conditions, $(C_A - C_{A,s}, T_r - T_{r,s}, T_j - T_{j,s}) = (1\text{kmol m}^{-3}, 30K, 40\text{K})$. Practically, since the optimization problem is solved in every sampling period Δ, the computational time for a single iteration cannot exceed Δ. As Δ is chosen to be 10 s in this application, for a feasible N, no iteration time can exceed 10 s. From Fig. 11, it is observed that, for the LMPC utilizing the first-principles slow model, the maximum allowable prediction horizon length is $N = 16$. Numerically, it is confirmed that, in this case, the longest iteration times when $N = 16$ and $N = 17$ are 9.59 s and 10.85 s, respectively. In contrast, for the LMPC with the SI slow model, the maximum allowable prediction horizon length is $N = 24$ as seen in Fig. 12. For this controller, the longest iteration times for prediction horizon lengths of $N = 24$ and $N = 25$ were calculated to be 9.83 s and 13.6 s, respectively. Finally, the time-integral of the LMPC cost function $\int_{t=0}^{t_p} L(x(\tau), u(\tau))\, \mathrm{d}\tau$ is calculated over the simulation period $t_p = 1$hr for the controller utilizing the first-principles model with $N = 16$ and the controller incorporating the SI slow model with $N = 24$. It is found that the costs are $L = 243.8$ and $L = 225.9$ for the first-principles model-based controller and SI-based controller, respectively. This implies that the LMPC based on the SI slow model, when maximizing the prediction horizon length to $N = 24$, outperforms the LMPC based on the first-principles slow model with its prediction horizon length maximized at $N = 16$, in terms of lower energy and faster convergence to the origin. This fact can also be observed from the state and input profiles, shown in Figs. 14 and 13, respectively. The LMPC based on the SI slow model takes more aggressive control action in the earlier MPC steps as seen in Fig. 13, causing the state to converge slightly faster to the origin as depicted in Fig. 14. The maximum allowable prediction horizon lengths and their corresponding costs are summarized in Table 2.

7. Conclusion

This article focused on the design of a Lyapunov-based MPC for a class of nonlinear singularly perturbed systems using only measurement data from processes. In singularly perturbed systems, due to the presence of time-scale multiplicities, a direct application of MPC without accounting for the evolution of the states in different time scales can lead to closed-loop performance deterioration or even closed-loop instability due to controller ill-conditioning. Hence, we proposed a method to first separate the slow and fast variables in the system and then design the MPC based on the reduced-order slow subsystem. Furthermore, due to the lack of a first-principles model in most practical applications, our method used only sampled experimental/industrial simulation data to reconstruct the reduced slow subsystem via a machine-learning method known as sparse identification. Subsequently, the theory was developed by deriving sufficient conditions for closed-loop stability under sample-and-hold implementation. Finally, the proposed LMPC design was applied to a non-isothermal reactor that exhibited time-scale separation. It was observed that the controllers yielded nearly identical performance for the same controller parameters. However, the LMPC based on the sparse-identified slow subsystem could implement superior controller parameters, such as a longer prediction horizon, due to its reduced complexity and, hence, lower computational time. As a result, the SI based LMPC outperformed the LMPC utilizing the first-principles model when the superior parameters were used for the former controller, demonstrating the practicality and benefits of designing MPC by reconstructing the reduced slow subsystem from measurement data.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CRediT authorship contribution statement

Fahim Abdullah: Conceptualization, Methodology, Software, Writing - original draft. **Zhe Wu:** Conceptualization, Methodology, Software, Writing - original draft. **Panagiotis D. Christofides:** Writing - review & editing.

Acknowledgments

Financial support from the National Science Foundation and the Department of Energy is gratefully acknowledged.

References

Abdullah, F., Wu, Z., Christofides, P.D., 2021. Data-based reduced-order modeling of nonlinear two-time-scale processes. Chem. Eng. Res. Des. 166, 1–9.
Alanqar, A., Durand, H., Christofides, P.D., 2015. On identification of well-conditioned nonlinear systems: application to economic model predictive control of nonlinear processes. AIChE J. 61 (10), 3353–3373.
Arnaldo, I., O'Reilly, U.-M., Veeramachaneni, K., 2015. Building predictive models via feature synthesis. In: Proceedings of the 2015 Annual Conference on Genetic and Evolutionary Computation. Association for Computing Machinery, New York, NY, USA, pp. 983–990.
Bai, Z., Wimalajeewa, T., Berger, Z., Wang, G., Glauser, M., Varshney, P.K., 2015. Low-dimensional approach for reconstruction of airfoil data via compressive sensing. AIAA Journal 53 (4), 920–933.
Brunton, S.L., Proctor, J.L., Kutz, J.N., 2016, Vol. 49, pp. 710–715. 10th IFAC Symposium on Nonlinear Control Systems NOLCOS 2016
Brunton, S.L., Proctor, J.L., Kutz, J.N., 2016. Discovering governing equations from data by sparse identification of nonlinear dynamical systems. Proceedings of the National Academy of Sciences 113 (15), 3932–3937.
Brunton, S.L., Tu, J.H., Bright, I., Kutz, J.N., 2014. Compressive sensing and low-rank libraries for classification of bifurcation regimes in nonlinear dynamical systems. SIAM J Appl Dyn Syst 13 (4), 1716–1732.
Champion, K.P., Brunton, S.L., Kutz, J.N., 2019. Discovery of nonlinear multiscale systems: sampling strategies and embeddings. SIAM J Appl Dyn Syst 18 (1), 312–333.
Chartrand, R., 2011. Numerical differentiation of noisy, nonsmooth data. ISRN Applied Mathematics 2011, 1–11.
Chen, X., Heidarinejad, M., Liu, J., de la Peña, D.M., Christofides, P.D., 2011. Model predictive control of nonlinear singularly perturbed systems: application to a large-scale process network. J Process Control 21 (9), 1296–1305.
Christofides, P.D., Teel, A.R., 1996. Singular perturbations and input-to-state stability. IEEE Trans Automat Contr 41 (11), 1645–1650.
Ellis, M., Heidarinejad, M., Christofides, P.D., 2013. Economic model predictive control of nonlinear singularly perturbed systems. J Process Control 23, 743–754.
Kokotović, P., Khalil, H.K., O'Reilly, J., 1986. Singular perturbation methods in control: Analysis and design. Academic Press, London.
Kumar, A., Christofides, P.D., Daoutidis, P., 1998. Singular perturbation modeling of nonlinear processes with nonexplicit time-scale multiplicity. Chem Eng Sci 53, 1491–1504.
Lin, Y., Sontag, E.D., 1991. A universal formula for stabilization with bounded controls. Systems & Control Letters 16, 393–397.
Mackey, A., Schaeffer, H., Osher, S., 2014. On the compressive spectral method. Multiscale Modeling & Simulation 12 (4), 1800–1827.
Ozolinš, V., Lai, R., Caflisch, R., Osher, S., 2013. Compressed modes for variational problems in mathematics and physics. Proceedings of the National Academy of Sciences 110 (46), 18368–18373.
Proctor, J.L., Brunton, S.L., Brunton, B.W., Kutz, J.N., 2014. Exploiting sparsity and equation-free architectures in complex systems. The European Physical Journal Special Topics 223 (13), 2665–2684.

Ricardez Sandoval, L., Budman, H., Douglas, P., 2008. Simultaneous design and control of processes under uncertainty: a robust modelling approach. J Process Control 18 (7), 735–752.
Ricardez-Sandoval, L.A., Budman, H.M., Douglas, P.L., 2009. Application of robust control tools to the simultaneous design and control of dynamic systems. Industrial & Engineering Chemistry Research 48 (2), 801–813.
Sanchez-Sanchez, K.B., Ricardez-Sandoval, L.A., 2013. Simultaneous design and control under uncertainty using model predictive control. Industrial & Engineering Chemistry Research 52 (13), 4815–4833.
Savitzky, A., Golay, M.J., 1964. Smoothing and differentiation of data by simplified least squares procedures.. Anal. Chem. 36, 1627–1639.
Schaeffer, H., Caflisch, R., Hauck, C.D., Osher, S., 2013. Sparse dynamics for partial differential equations. Proceedings of the National Academy of Sciences 110 (17), 6634–6639.
Wächter, A., Biegler, L.T., 2006. On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming. Math Program 106, 25–57.
Wang, W.-X., Yang, R., Lai, Y.-C., Kovanis, V., Grebogi, C., 2011. Predicting catastrophes in nonlinear dynamical systems by compressive sensing. Phys. Rev. Lett. 106 (15), 154101.
Wu, Z., Tran, A., Rincon, D., Christofides, P.D., 2019. Machine learning-based predictive control of nonlinear processes. part i: theory. AIChE J. 65, e16729.
Wu, Z., Tran, A., Rincon, D., Christofides, P.D., 2019. Machine-learning-based predictive control of nonlinear processes. part ii: computational implementation. AIChE J. 65, e16734.