Error-Triggered On-Line Model Identification for Model-Based Feedback Control

Anas Alanqar and Helen Durand
Dept. of Chemical and Biomolecular Engineering, University of California, Los Angeles, CA 90095

Panagiotis D. Christofides
Dept. of Chemical and Biomolecular Engineering and Dept. of Electrical Engineering, University of California, Los Angeles, CA 90095

DOI 10.1002/aic.15430
Published online August 3, 2016 in Wiley Online Library (wileyonlinelibrary.com)

In industry, it may be difficult in many applications to obtain a first-principles model of the process, in which case a linear empirical model constructed using process data may be used in the design of a feedback controller. However, linear empirical models may not capture the nonlinear dynamics over a wide region of state-space and may also perform poorly when significant plant variations and disturbances occur. In the present work, an error-triggered on-line model identification approach is introduced for closed-loop systems under model-based feedback control strategies. The linear models are re-identified on-line when significant prediction errors occur. A moving horizon error detector is used to quantify the model accuracy and to trigger the model re-identification on-line when necessary. The proposed approach is demonstrated through two chemical process examples using a model-based feedback control strategy termed Lyapunov-based economic model predictive control (LEMPC). The chemical process examples illustrate that the proposed error-triggered on-line model identification strategy can be used to obtain more accurate state predictions to improve process economics while maintaining closed-loop stability of the process under LEMPC. © 2016 American Institute of Chemical Engineers *AIChE J*, 63: 949–966, 2017
Keywords: nonlinear systems, on-line model identification, economic model predictive control, process control, process optimization, process economics, nonlinear processes

Introduction

The operational excellence and energy management of chemical and petrochemical processes rely on finding industrial solutions for the global energy demand, which led engineers to develop technologies that promote optimal process operation. To achieve improvements in process control, model-based control strategies, such as Lyapunov-based control and model predictive control (MPC), have been introduced. These types of controllers can improve process operation and thus may increase profit. A fairly recent model-based control strategy termed economic model predictive control (EMPC), for example, performs dynamic economic optimization incorporating predictions of future process states and state or output feedback to establish optimal time-varying operation under constraints.[1–4] The potential of model-based control strategies to improve process efficiency and to obtain desired closed-loop response characteristics makes these strategies desirable for use in industry.

A dynamic process model is required for any process for which model-based control is proposed and such models can be established through first-principles or empirical modeling.[5] First-principles models mathematically describe observed phenomena; the development of such models is difficult for processes that are complex and/or poorly understood. Numerous research efforts have been dedicated to the development of highly reliable model identification methods that require only input and output data to develop linear and nonlinear empirical models that can be used when first-principles models are undeveloped or impractical for on-line process control computations.[5–7]

A well-known class of empirical models are those designed using subspace model identification (SMI) methods, which are state-space model identification techniques for multiple-input multiple-output (MIMO) systems.[7–11] SMI methods are non-iterative robust methods that take into account multivariable interactions and result in highly reliable models.[7,12] Linear subspace identification algorithms include the Canonical Variate Algorithm (CVA),[13] the multivariable output error state-space (MOESP) algorithm,[10,11] and numerical algorithms for subspace state-space system identification (N4SID).[12] These methods have been successful in industrial applications and they result in numerically stable models.[6,8,10,14] SMI methods have also been considered for use in model predictive control (MPC) and EMPC.[9,15] Recursive subspace identification, in which the identified model is updated in order to correct for disturbances and nonlinearities, has also been an active area of research.[16] Most of this research has focused on the mathematical analysis of the methods used to update the identified models, including recursive least-squares, least mean squares, and recursively updating the singular value decomposition.[16–18]

Correspondence concerning this article should be addressed to P. D. Christofides at pdc@seas.ucla.edu.

However, such methods have not focused on determining an approach for triggering the re-identification of a model when necessary.

In this work, an error-triggered on-line model identification approach is introduced for model-based control strategies. The re-identification of the model is conducted on-line to reduce plant-model mismatch that occurs very often in the chemical industry due to various reasons such as significant disturbances, catalyst deactivation in reactors, and upsets in feed streams, to name a few. Also, the on-line model identification method can be used to update an empirical model when significant plant-model mismatch is detected because the region of operation shifts and the current model no longer captures the nonlinear dynamics. A moving horizon error detector monitors the prediction error between the states predicted by the empirical model and the measured states of the process. When the error exceeds a pre-specified threshold, the detector triggers an on-line model re-identification which is performed using the most recently generated input/output data. The approach is applied in the context of Lyapunov-based economic model predictive control (LEMPC) for nonlinear process systems through two chemical process examples. The first example demonstrates the ability of the proposed approach to improve the accuracy of the predicted states when significant plant-model mismatch occurs due to variations in the plant (catalyst deactivation). In the second example, the operating region is expanded gradually to allow the process to operate in a larger region for improved profit, and the proposed approach improves the accuracy of states predicted by the LEMPC over a larger region of state-space.

Preliminaries

Notation

The symbol x^T is used to denote the transpose of the vector x. The operator $|\cdot|$ designates the 2-norm of a vector. A continuous function $\alpha : [0, a) \to [0, \infty)$ is said to belong to class $\mathcal{K}$ if it is strictly increasing and equal to zero only when evaluated at zero. The symbol Ω_ρ is used to denote a level set of a sufficiently smooth scalar function $V(x)$ ($\Omega_\rho := \{x \in R^n : V(x) \leq \rho\}$). A square diagonal matrix is designated as diag(v) where the diagonal elements equal the components of the vector v. The symbol $\Delta > 0$ denotes the sampling period. The notation $S(\Delta)$ signifies the set of piecewise-constant vector functions with period Δ.

Class of systems

The class of nonlinear systems considered in this work is of the following form

$$\frac{dx(t)}{dt} = f(x(t), u(t), w(t)) \tag{1}$$

where $x \in R^n$, $u \in R^m$, and $w \in R^l$ are the system state vector, the manipulated input vector, and the disturbance vector respectively. Physical limitations on the available control energy set by actuator constraints are considered by restricting the control actions to the convex set $U := \{u \in R^m : u_i^{\min} \leq u_i \leq u_i^{\max},\ i=1, \ldots, m\}$. We consider a bounded disturbance vector in this work (i.e., $w \in W := \{w\ :\ |w(t)| \leq \theta\ \forall\ t\}$). Measurements of the entire state vector $x(t_k)$ are assumed to be available at each sampling time $t_k = k\Delta$, $k=0, 1, \ldots$.

We restrict our discussion to the class of stabilizable nonlinear systems for which there exists a controller $h(x) \in U$ that can render the origin of the nominal ($w(t) \equiv 0$) closed-loop system of Eq. 1 asymptotically stable in the sense that there exists a sufficiently smooth Lyapunov function $V : R^n \to R_+$ that satisfies the following inequalities[19,20]

$$\alpha_1(|x|) \leq V(x) \leq \alpha_2(|x|) \tag{2a}$$

$$\frac{\partial V(x)}{\partial x} f(x, h(x), 0) \leq -\alpha_3(|x|) \tag{2b}$$

$$\left| \frac{\partial V(x)}{\partial x} \right| \leq \alpha_4(|x|) \tag{2c}$$

for all x in an open neighborhood $D \subseteq R^n$ that includes the origin and $\alpha_j(\cdot)$, $j=1,\ 2,\ 3,\ 4$, are class $\mathcal{K}$ functions. Various stabilizing controllers that take into account input constraints have been developed for several classes of nonlinear systems.[21–23] The stability region of the closed-loop system is taken to be a level set $\Omega_\rho \subset D$ where the time derivative of the Lyapunov function is strictly negative ($(dV/dt) < 0$). In addition, the origin of the system of Eq. 1 is rendered practically stable[24] when the controller $h(x)$ is applied in a sample-and-hold fashion for a sufficiently small sampling period. The function f is assumed to be locally Lipschitz on $\Omega_\rho \times U \times W$ and the origin is taken to be an equilibrium of the unforced system of Eq. 1 (i.e., $f(0, 0, 0) = 0$).

The proposed on-line model identification approach develops linear empirical models to predict the evolution of the state of the system of Eq. 1. Although the method discussed in this work extends to a wide range of linear empirical models such as input/output models, we will demonstrate the scheme where the empirical models obtained on-line are state-space linear time-invariant (LTI) models of the form

$$\frac{dx(t)}{dt} = A_i x(t) + B_i u(t) \tag{3}$$

where A_i and B_i are constant matrices of appropriate dimensions corresponding to the $i-th$ model identification performed ($i=1,\ \ldots,\ \tilde{M}$).

We assume that a set of stabilizing controllers $h_{L1}(x)$, $h_{L2}(x)$, $\ldots$, $h_{L\tilde{M}}(x)$ designed based on the empirical models exists such that each controller renders the origin of the closed-loop system of Eq. 1 asymptotically stable and yields a sufficiently smooth Lyapunov function $\hat{V} : R^n \to R_+$ with the following properties[19]

$$\hat{\alpha}_1(|x|) \leq \hat{V}(x) \leq \hat{\alpha}_2(|x|) \tag{4a}$$

$$\frac{\partial \hat{V}(x)}{\partial x} f(x, h_{Li}(x), 0) \leq -\hat{\alpha}_{3_i}(|x|),\ i=1,\ \ldots,\ \tilde{M} \tag{4b}$$

$$\left| \frac{\partial \hat{V}(x)}{\partial x} \right| \leq \hat{\alpha}_4(|x|) \tag{4c}$$

for all $x \in D_{Li} \subseteq R^n$ where D_{Li} is an open neighborhood that includes the origin and the functions $\hat{\alpha}_j(\cdot)$, $j=1,\ 2,\ 4$, and $\hat{\alpha}_{3_i}$, $i=1,\ \ldots,\ \tilde{M}$, are class $\mathcal{K}$ functions. The system of Eq. 1 under the controller $h_{Li}(x)$ has the stability region $\Omega_{\hat{\rho}_i} \subset D_{Li}$, $i=1,\ \ldots,\ \tilde{M}$.

Lyapunov-based economic model predictive control

The model-based controller that will be used in the chemical process examples presented in this work is Lyapunov-based economic model predictive control (LEMPC).[3] LEMPC uses a receding horizon control strategy that minimizes a cost function while incorporating stability constraints based on the

explicit stabilizing controller $h(x)$ in its design. The formulation of LEMPC is

$$\min_{u\in S(\Delta)} \int_{t_k}^{t_{k+N}} L_e(\tilde{x}(\tau),u(\tau))\ d\tau \tag{5a}$$

$$\text{s.t.} \quad \dot{\tilde{x}}(t)=f(\tilde{x}(t),u(t),0) \tag{5b}$$

$$\tilde{x}(t_k)=x(t_k) \tag{5c}$$

$$u(t)\in U,\ \forall\ t\in[t_k,t_{k+N}) \tag{5d}$$

$$V(\tilde{x}(t))\le\rho_e,\ \forall\ t\in[t_k,t_{k+N}) \quad \text{if } x(t_k)\in\Omega_{\rho_e} \tag{5e}$$

$$\frac{\partial V(x(t_k))}{\partial x}f(x(t_k),u(t_k),0) \le \frac{\partial V(x(t_k))}{\partial x}f(x(t_k),h(x(t_k)),0) \quad \text{if } x(t_k)\notin\Omega_{\rho_e} \tag{5f}$$

where the optimization variable is the control input trajectory over the prediction horizon $N\Delta$. LEMPC uses the process dynamic model (Eq. 5b) to predict the state trajectory $\tilde{x}(t)$ over time starting from the initial condition in Eq. 5c which is obtained from the state measurement at time t_k. The LEMPC design takes into account constraints on the manipulated inputs in Eq. 5d. The stage cost L_e (Eq. 5a) is formulated to represent the process economics.

Based on the state measurement of Eq. 5c, either the Mode 1 (Eq. 5e) or the Mode 2 (Eq. 5f) stability constraint is active. In Mode 1, time-varying operation is promoted to maximize profit while maintaining the state in a region $\Omega_{\rho_e}\subset\Omega_\rho$ chosen to make Ω_ρ invariant in the presence of disturbances. If the measured process state is outside of Ω_{ρ_e}, Mode 2 is activated to compute control actions that decrease the Lyapunov function value and force the state back into Ω_{ρ_e}. The solution of the LEMPC optimization problem is denoted as $u^*(t|t_k)$, $t\in[t_k,t_{k+N})$, but only $u^*(t_k|t_k)$ is implemented at each sampling time.

Lyapunov-based economic model predictive control with empirical models

The plant model of Eq. 5b may be unavailable, in which case it can be replaced by an empirical model. In this work, we will replace the nonlinear plant model of Eq. 5b with the $i-th$ linear empirical model, $i=1,\ \ldots,\ \tilde{M}$ (which is the model last identified by the error-triggered on-line model identification procedure prior to the sampling time t_k). The empirical model is also used to design $h_{Li}(x)$ and $\hat{V}(x)$ for the Lyapunov-based constraints in Eq. 5. The level set $\Omega_{\hat{\rho}_{ei}}\subset\Omega_{\hat{\rho}_i}$ that prompts the switch between Mode 1 and Mode 2 is chosen such that the controller maintains operation of the process of Eq. 1 within $\Omega_{\hat{\rho}_i}$ in the presence of bounded disturbances. The formulation of LEMPC using the $i-th$ linear empirical model is[15]

$$\min_{u\in S(\Delta)} \int_{t_k}^{t_{k+N}} L_e(\hat{x}(\tau),u(\tau))\ d\tau \tag{6a}$$

$$\text{s.t.} \quad \dot{\hat{x}}(t)=A_i\hat{x}(t)+B_iu(t) \tag{6b}$$

$$\hat{x}(t_k)=x(t_k) \tag{6c}$$

$$u(t)\in U,\ \forall\ t\in[t_k,t_{k+N}) \tag{6d}$$

$$\hat{V}(\hat{x}(t))\le\hat{\rho}_{ei},\ \forall\ t\in[t_k,t_{k+N}) \quad \text{if } x(t_k)\in\Omega_{\hat{\rho}_{ei}}, \tag{6e}$$

$$\frac{\partial\hat{V}(x(t_k))}{\partial x}(A_ix(t_k)+B_iu(t_k)) \le \frac{\partial\hat{V}(x(t_k))}{\partial x}(A_ix(t_k)+B_ih_{Li}(x(t_k))) \quad \text{if } x(t_k)\notin\Omega_{\hat{\rho}_{ei}} \tag{6f}$$

where the notation follows that in Eq. 5 except that $\hat{x}(t)$ is the predicted state of the system using the linear empirical model (Eq. 6b), starting from a measurement of the actual process state (Eq. 6c) to predict the evolution of the system of Eq. 1.

Remark 1. In the formulation of Eq. 6, $\hat{V}$ is not updated when the model of Eq. 6b is updated. Although it may be replaced with $\hat{V}_i$, this is not required if the stability region of an updated model can be found to be a level set of the same Lyapunov function as was used for the prior model.

Remark 2. A key feature of the LEMPCs in Eqs. 5 and 6 is that they may not drive the process to a steady-state, but rather operate it in a time-varying fashion within a stability region, when the cost function does not have its minimum at a steady-state. When such dynamic operation is achieved, the time-varying nature of the input trajectories generated using LEMPC can result in persistent excitation of the process states, which makes the inputs ideal for on-line model identification. The chemical process examples in this work demonstrate the time-varying nature of inputs that may be calculated by an LEMPC.

Error-Triggered On-Line Model Identification

This section discusses the proposed error-triggered on-line model identification method.

Error-triggering mechanism for on-line model identification

In this section, we describe an error-triggering mechanism that can trigger on-line updates of the model utilized in the design and implementation of a model-based controller. A major advantage of this mechanism is that it can prevent constant updating of the process model, which may be computationally expensive and result in frequent changes to the control law that are undesirable. It also prevents the use of an inaccurate model when there is significant plant-model mismatch.

In the proposed method, a moving horizon error detector quantifies the accuracy of an empirical model by calculating the following moving horizon error metric e_d at times t_k

$$e_d(t_k)=\sum_{r=0}^{M}\sum_{j=1}^{n}\frac{|x_{p,j}(t_{k-r})-x_j(t_{k-r})|}{|x_j(t_{k-r})|} \tag{7}$$

where M is the number of sampling periods before t_k that contribute to the quantification of the prediction error, $x_j(t_{k-r})$, $r=0,\ \ldots,\ M$, are the past measurements of the process states at sampling periods between t_{k-M} and t_k, and $x_{p,j}(t_{k-r})$, $r=0,\ \ldots,\ M$, are the predictions of the past states of the system from a linear empirical model. A threshold value $e_{d,T}$ is set for the error metric, and when the moving horizon error detector determines that this threshold has been exceeded, it triggers model re-identification. There are several parameters that need to be defined to carry out this error-triggered approach: the number of input and output data points

N_d that must be kept for model identification when it is triggered, the length M of the moving horizon used in the calculation of e_d, and the threshold $e_{d,T}$ of the error above which model re-identification is triggered. To determine these parameters initially, the following strategy is proposed:

Step 1. When no initial linear model is available (e.g., through the linearization of a first-principles model), the process is initially excited off-line in open-loop using long sequences of standard input types (e.g., impulse or step inputs) to excite the important dynamics. The corresponding output data is collected. System identification is carried out on the (open-loop) input/output data using standard techniques (e.g., determination of a large order model followed by model order reduction[7]) to determine an $i = 1$ empirical model that captures the dominant process dynamics. The number of input/output data points N_d that need to be stored for a possible future system identification can be set to the number of input/output data points required to identify the $i = 1$ empirical model.[7,12]

Step 2. The value of M to be used in the calculation of e_d must be long enough such that disturbances common during normal operation do not significantly affect e_d (which could lead to unnecessary error-triggering) but are smoothed out. However, M also should not be longer than necessary because this would require unnecessary data storage and processing. One method for determining M is by calculating the value of $e_d(t_k)$, $t_k > t_M$, at every sampling period for a set of input/output data collected during normal process operation (in closed-loop under the model-based controller) in the region of operation for which the $i = 1$ empirical model was developed and validated. This calculation is then repeated for various values of M. The minimum and maximum values of e_d for a given value of M may be significantly different if M is small, since then any disturbance or measurement noise within the moving horizon contributes significantly to the value of e_d. As M is increased, however, the effect of disturbances and measurement noise will become less significant. At some point, it would be expected that the minimum and maximum values of e_d will not change much when M is further increased; in this case, the smallest value of M for which the minimum and maximum values of e_d seem to have reached their approximate final value could be chosen to be used in Eq. 7. From this, it is seen that the statistical properties of $w(t)$ will affect the value of M for a given process.

Step 3. The value of $e_{d,T}$ is determined off-line, based on the chosen value of M, such that measurement noise, small constant disturbances, and time-varying disturbances that cause reasonably accurate predictions with the current model do not trigger model re-identification. One method for achieving this is by analyzing the statistical properties of e_d for a set of closed-loop input/output data corresponding to normal process operation in the region within which the linear empirical model was developed and validated. For example, the maximum value of e_d for such data could be determined with the selected value of M, and then the threshold could be chosen to be a reasonable percentage greater than the maximum value of e_d observed for this normal operating data (which should include the disturbances and measurement noise that regularly affect the system). Other statistical measures (e.g., choosing the value of $e_{d,T}$ to be several standard deviations above its mean value from the normal operating data) could also be used, and the appropriate measure to use will depend on the system analyzed. It is noted that even if there were no disturbances or measurement noise, e_d would be expected to have a value because the linear empirical model identified is unlikely to fully capture the nonlinear dynamics of a given process.

The lack of a formula for obtaining N_d, M, and $e_{d,T}$ does not pose practical limitations because, as will be discussed further in later sections, the on-line, data-based methodology employed allows for constant monitoring of the process performance under a controller based on the current linear empirical model such that "poor" choices of the parameters can be detected and the parameters adjusted.

Remark 3. For consistency with Step 1 as presented above, in the remainder of the manuscript we will assume that no first-principles process model is available and we will call the linear models developed before the initiation of error-triggered on-line model identification the initial linear empirical models. However, it is not required that these models be empirical. The initial linear model obtained in Step 1 of the above procedure may be obtained using a linearization of a nonlinear first-principles model if such a model is available. In this case, the utility of on-line updating of the linear model using the error-triggered model identification method is that it prevents the need for using a nonlinear model and may aid in capturing the process dynamics better than a first-principles model as the process dynamics change in time.

Remark 4. The value of $x_{p,j}(t_{k-r})$ may be calculated in multiple ways. For example, one method is to calculate each value of $x_{p,j}(t_{k-r})$ by numerically integrating the linear empirical model of Eq. 3, starting from the measured state $x(t_{k-r-1})$ at the previous sampling time, and using the input applied for $t \in [t_{k-r-1}, t_{k-r})$ in the integration. Another method is to determine the value of each $x_{p,j}(t_{k-r})$ by integrating the linear model, starting from the measurement of the state $x(t_{k-M-1})$ and applying the sequence of control actions implemented on the process throughout time. If more than one model has been used between t_{k-M-1} and t_k, the applicable model should be used for the integration corresponding to its period of use.

Remark 5. Practically, when the model-based controller is updated due to the identification of a new model, it may be necessary to include additional precautions in the control system design. For example, a constraint or saturation could be imposed when the controller is initially updated that prevents the inputs calculated from the new controller from differing more than a certain amount from the control actions calculated most recently by the prior controller. If this is used, the closed-loop stability properties of the resulting controller should be considered.

Remark 6. In model-based controllers such as standard tracking model predictive control (MPC), the replacement of a prior linear empirical model in the MPC with a newer model obtained from the error-triggered on-line model identification strategy may be sufficient for keeping the control law up-to-date. Some model-based control laws may require further computation to determine the new control law after an updated linear empirical model is developed (e.g., Sontag's control law will need to be re-calculated based on the new model[23]) Other control laws may have additional aspects that require them to be further adjusted as the linear empirical models are updated. For example, LEMPC may require that the process model be updated and also that other

components of the control law in Eq. 6 be updated (e.g., h_{Li}).

Implementation strategy for error-triggered on-line model identification

Once the values of N_d, M and $e_{d,T}$ are set according to the methodology of the prior section, the proposed error-triggered on-line model identification strategy can be executed with the following implementation strategy:

Step 1. An initial linear empirical model of the plant (A_1 and B_1) is developed (this becomes the "current model" for the process). This model is used to design the model-based controller.

Step 2. The process is operated under the model-based controller designed based on the current linear empirical model, and input/output data (up to N_d values of each) are collected and stored for possible future model identification. The moving horizon error detector is initiated at t_M to calculate $e_d(t_k)$.

Step 3. As the current linear model begins to fail to describe the process dynamics (due to, e.g., variations in the plant or changes in the region of operation), $e_d(t_k)$ will increase and when $e_d(t_k)$ exceeds $e_{d,T}$, the most recent set of N_d values of input and output data (collected up to time t_k) are used to identify a new model on-line to become the current model for use in updating the model-based controller formulation.

Step 4. Steps 2–3 are repeated as process operation continues.

Applications of Error-Triggered On-Line Model Identification

In this section, we present two applications of the error-triggered on-line model identification methodology (Steps 1–4 of the prior section) to demonstrate the flexibility of the approach and its utility in a variety of circumstances.

Application of error-triggered on-line model identification to plant variations

The first application to be discussed is that in which the plant model changes in time. This may occur, for example, due to catalyst deactivation that affects the reaction rates in the process, due to heat exchanger fouling that affects the rate of heat transfer, or due to changes in the valve dynamics with time as dynamic valve nonlinearities such as stiction worsen in time due to valve degradation. For this application, it is possible that as the plant model changes in time, the values of N_d, M, and $e_{d,T}$ originally determined may no longer be the appropriate values. For example, N_d may need to be increased to obtain a more accurate model, and the combination of M and $e_{d,T}$ may cause the linear empirical model to be updated more or less frequently than necessary. This can be handled on-line by varying the parameters in small increments until better performance is obtained. For example, N_d may be lengthened until the linear empirical models are shown to better capture the nonlinear dynamics in the current region of state-space. M and $e_{d,T}$ may be increased if the on-line model updates are triggered too frequently even when the linear empirical models being identified do not result in poor closed-loop performance, or they may be decreased if the model updates are not triggered frequently enough and process and controller performance degrades. Alternatively, experiments can be performed to determine new values of these parameters.

Remark. If the values of N_d, M, and $e_{d,T}$ are picked appropriately such that the model-based controller used at any given time is based on a reasonably accurate empirical model, this controller may often be stabilizing if the computation time required by the moving horizon error detector and for model updates is short compared to the process dynamics (this prevents the process state from changing significantly before errors are detected and the model is updated). To develop a rigorous proof of closed-loop stability of a process under the error-triggered on-line model identification procedure as plant variations occur in time, however, a number of mathematical assumptions would need to be made that are specific to each model-based controller type. Proofs of closed-loop stability of processes under various model-based controllers and feasibility of such controllers as applicable are addressed in many works (e.g.,[3,15,19,25,26] and the references therein), and similar methods could be investigated for nonlinear processes under model-based controllers designed based on linear empirical models developed from the error-based triggering approach. An example of an assumption that may be considered in the proofs of feasibility and closed-loop stability of a process under the LEMPC of Eq. 6 is that the closed-loop state is within both $\Omega_{\hat{\rho}_{i-1}}$ and $\Omega_{\hat{\rho}_i}$ at the time that the model used in the design of the LEMPC of Eq. 6 changes from the $(i-1)-th$ to the $i-th$ linear empirical model. The primary purpose of the present paper is not for providing rigorous closed-loop stability proofs (although the more accurate empirical models may aid closed-loop system stabilizability), but is for providing more accurate models for model-based control strategies so that more desirable control actions (e.g., more economically beneficial or more capable of meeting process constraints) can be calculated.

Application of error-triggered on-line model identification to plant variations: application to a chemical process example

In this section, we demonstrate the proposed error-triggered on-line model identification procedure for the control of a benchmark chemical reactor for which plant model changes occur, specifically catalyst deactivation. The catalytic oxidation of ethylene (C_2H_4) in a non-isothermal continuous stirred tank reactor (CSTR) is considered. The ethylene is oxidized with air to produce the desired ethylene oxide (C_2H_4O) product. Two combustion reactions that consume ethylene oxide and ethylene occur concurrently in the reactor as presented in the following chemical reactions

$$C_2H_4+\frac{1}{2}O_2\xrightarrow{R_1}C_2H_4O \tag{R1}$$

$$C_2H_4+3O_2\xrightarrow{R_2}2CO_2+2H_2O \tag{R2}$$

$$C_2H_4O+\frac{5}{2}O_2\xrightarrow{R_3}2CO_2+2H_2O \tag{R3}$$

The rates of the reactions R_1, R_2, and R_3 are given by the following rate laws[27]

$$R_1=k_1\exp\left(\frac{-E_1}{RT}\right)P_E^{0.5} \tag{8a}$$

$$R_2=k_2\exp\left(\frac{-E_2}{RT}\right)P_E^{0.25} \tag{8b}$$

$$R_3 = k_3 \exp\left(\frac{-E_3}{RT}\right) P_{EO}^{0.5} \tag{8c}$$

where the pre-exponential factors are k_1, k_2, and k_3 and the activation energies are E_1, E_2, and E_3. T is the temperature and R is the gas constant. P_E and P_{EO} are the partial pressures of ethylene (P_E) and of ethylene oxide (P_{EO}) and it is assumed that the gas mixture in the reactor is an ideal gas, and thus, the partial pressures can be written in terms of the molar concentrations. The dimensionless first-principles dynamic model which is derived from mass and energy balances for this process from[28] is of the following form

$$\frac{dx_1(t)}{dt} = u_1(1 - x_1 x_4) \tag{9a}$$

$$\frac{dx_2(t)}{dt} = u_1(u_2 - x_2 x_4) - A_1 e^{\frac{\gamma_1}{x_4}} (x_2 x_4)^{0.5} - A_2 e^{\frac{\gamma_2}{x_4}} (x_2 x_4)^{0.25} \tag{9b}$$

$$\frac{dx_3(t)}{dt} = -u_1 x_3 x_4 + A_1 e^{\frac{\gamma_1}{x_4}} (x_2 x_4)^{0.5} - A_3 e^{\frac{\gamma_3}{x_4}} (x_3 x_4)^{0.5} \tag{9c}$$

$$\frac{dx_4(t)}{dt} = \frac{u_1}{x_1}(1 - x_4) + \frac{B_1}{x_1} e^{\frac{\gamma_1}{x_4}} (x_2 x_4)^{0.5} + \frac{B_2}{x_1} e^{\frac{\gamma_2}{x_4}} (x_2 x_4)^{0.25} + \frac{B_3}{x_1} e^{\frac{\gamma_3}{x_4}} (x_3 x_4)^{0.5} - \frac{B_4}{x_1}(x_4 - u_3) \tag{9d}$$

where the dimensionless variables x_1, x_2, x_3, and x_4 correspond to the gas density in the reactor, ethylene concentration, ethylene oxide concentration, and reactor temperature, respectively. The manipulated inputs u_1, u_2, and u_3 are the dimensionless feed volumetric flow rate, ethylene concentration of the reactor feed, and coolant temperature, respectively. The manipulated inputs are bounded by physical limitations on actuators and hence, the inputs are constrained to belong to the following convex sets: $0.0704 \leq u_1 \leq 0.7042$, $0.2465 \leq u_2 \leq 2.4648$, $0.6 \leq u_3 \leq 1.1$. The values of the parameters of this model are presented in Table 1. The CSTR has an asymptotically stable steady-state that occurs at $x_s^T = [x_{1s}\ x_{2s}\ x_{3s}\ x_{4s}] = [0.998\ 0.424\ 0.032\ 1.002]$ when $[u_{1s}\ u_{2s}\ u_{3s}] = [0.35\ 0.5\ 1.0]$.

The CSTR is controlled by an LEMPC with the goal of feeding the ethylene to the reactor in a manner that maximizes the average yield of ethylene oxide. The average yield of ethylene oxide, which quantifies the amount of ethylene oxide produced compared to the amount of ethylene fed to the reactor, over a time period from t_0 to t_e, is given by

$$Y(t_e) = \frac{\int_{t_0}^{t_e} u_1(\tau) x_3(\tau) x_4(\tau)\ d\tau}{\int_{t_0}^{t_e} u_1(\tau) u_2(\tau)\ d\tau} \tag{10}$$

where t_e is an integer multiple of the length t_f of an operating period. Because the amount of reactant material available is fixed, the time-averaged molar flow rate of ethylene that can be fed to the reactor in an operating period is limited by the following constraint

$$\frac{1}{t_f} \int_{(j-1)t_f}^{jt_f} u_1(\tau) u_2(\tau)\ d\tau = u_{1s} u_{2s} = 0.175 \tag{11}$$

where j is the operating period number ($j = 1, 2, \ldots$). This constraint ensures that in each operating period, the amount of ethylene fed to the reactor is the same as that which would have been fed under steady-state operation. Because the integral input constraint of Eq. 11 fixes the value of the denominator in Eq. 10, the LEMPC seeks to maximize the following function

$$\int_{t_0}^{t_e} L_e(x, u) = \int_{t_0}^{t_e} u_1(\tau) x_3(\tau) x_4(\tau) d\tau \tag{12}$$

By maximizing this objective, the ethylene oxide yield is maximized subject to the integral material constraint that is enforced due to restrictions on the available feedstock.

Table 1. Dimensionless Parameters of the Ethylene Oxidation CSTR

$A_1 = 92.8$	$B_2 = 10.39$	$\gamma_2 = -7.12$
$A_2 = 12.66$	$B_3 = 2170.57$	$\gamma_3 = -11.07$
$A_3 = 2412.71$	$B_4 = 7.02$	
$B_1 = 7.32$	$\gamma_1 = -8.13$	

The first-principles nonlinear process model in Eq. 9 is assumed to be unavailable from a controller design point of view such that an empirical model must be obtained to formulate an LEMPC that meets the above objective and constraints. To construct such an empirical model that captures the dynamics within a region around the CSTR steady-state, a large number of input step changes of varying magnitudes were applied to the CSTR from the steady-state and the corresponding output data was collected. The ordinary multivariable output error state-space (MOESP)[10] algorithm was applied to this data to obtain the initial ($i = 1$) linear state-space empirical model for the CSTR of Eq. 9. This model was validated using a wide range of step, impulse, and sinusoidal input responses and is expressed by the following matrices

$$A_1 = \begin{bmatrix} -0.349 & 0.00051 & 0.00825 & -0.349 \\ -0.00488 & -0.374 & 0.0374 & -0.369 \\ 0.00109 & 0.0213 & -0.452 & 0.0653 \\ -0.0078 & 0.0259 & 0.0204 & -7.24 \end{bmatrix}$$

$$B_1 = \begin{bmatrix} -0.00011 & -0.000149 & -0.0239 \\ 0.0757 & 0.349 & -0.0194 \\ -0.0315 & 0.000208 & 0.00426 \\ -0.0173 & -0.00264 & 6.529 \end{bmatrix} \tag{13}$$

Since it is assumed that the only model available for controller design is the empirical model, the model of Eq. 13 is used to design the Lyapunov-based controller used in LEMPC. The controller can be represented as a vector with three components: $h_{L1}^T(x) = [h_{L1,1}(x)\ h_{L1,2}(x)\ h_{L1,3}(x)]$, where $h_{L1}(x_s) = 0$. The control laws $h_{L1,1}(x)$ and $h_{L1,2}(x)$ were set to the steady-state values of u_1 and u_2 to meet the material constraint of Eq. 11. The control law $h_{L1,3}(x)$ was designed using the standard linear quadratic regulator (LQR) with a quadratic objective defined using the A_1 matrix and the third column of the B_1 matrix (Eq. 13) and taking both the Q and R weighting matrices to be the identity matrix. This results in the control law $u_3 = h_{L1,3}(x) = -K(x - x_s) + u_{3s}$, with K equal to $[-0.287\ -0.276\ 0.023\ 0.405]$. A quadratic Lyapunov function of the form $\hat{V}(x) = (x - x_s)^T P (x - x_s)$ was utilized to characterize the stability region of the closed-loop system of Eq. 9 under the stabilizing controller h_{L1} with the positive definite matrix

P defined as $P = \text{diag}[20\ 30\ 40\ 10]$. Extensive simulations of the closed-loop system under the Lyapunov-based controller $h_{L1}(x)$ were conducted to define the level set $\Omega_{\hat{\rho}_{e1}}$ with $\hat{\rho}_{e1}=87.4$. This is a region within which the nonlinear dynamics of Eq. 9 are well-captured by the linear model of Eq. 13.

Although this chemical process example will be used to illustrate the error-triggered on-line model identification procedure in the presence of plant variations, we will first demonstrate that the initial linear empirical model performs well in the time period before any catalyst deactivation (plant variation) occurs. To demonstrate this, two LEMPC schemes, one of the form of Eq. 6 and the other of the form of Eq. 5, both with the additional material constraint of Eq. 11, were designed to compare closed-loop behavior. Both used the cost function of Eq. 12, the upper and lower bounds on u_1, u_2, and u_3 described above, and the same Lyapunov-based controller and stability region. The process model utilized in the first LEMPC was the empirical model of Eq. 13, while the second LEMPC utilized the first-principles model of Eq. 9. In all simulations for this example, the LEMPC designs had prediction horizons of $N = 10$, sampling periods of $\Delta=0.1$, and operating periods of 100 sampling periods ($t_f = 10$). The interior-point solver IPOPT[29] was used to solve the LEMPC optimization problems. The empirical LEMPC and the first-principles LEMPC were both applied to the CSTR model of Eq. 9. The reactor was initialized off steady-state at $x_I^T=[x_{1I}\ x_{2I}\ x_{3I}\ x_{4I}]=[0.997\ 1.264\ 0.209\ 1.004]$ and closed-loop simulations over 10 operating periods for each case were completed. The explicit Euler numerical integration method was used in all simulations for this example with an integration step size of $h=10^{-4}$. The closed-loop trajectories for the CSTR under both LEMPC schemes are shown in Figures 1 and 2 which demonstrate very similar behavior. The average yield of the first-principles LEMPC over 10 operating periods was 8.98, compared to 8.93 for the empirical LEMPC. The agreement between the trajectories and yields of the process under the first-principles and empirical LEMPCs demonstrates that the initial linear empirical model is capable of adequately describing the process behavior before plant variation occurs. It is noted that the periodic nature of the trajectories is consistent with prior literature for this example (e.g., Refs. 28 and 30) which demonstrated that time-varying operation can be economically beneficial for the ethylene oxide production process, and also other literature on optimal periodic operation (e.g., Refs. 31 and 32).

After the 10 operating periods, plant variations begin to occur (a reduction in the reaction pre-exponential factor is assumed to occur due to catalyst deactivation). Specifically, the pre-exponential factor values for reactions R_1, R_2, and R_3 are decreased by 40% gradually throughout nine operating periods. The pre-exponential factors are decreased by 10% of their original values at the beginning of the eleventh and twelfth operating periods, reaching the values $0.8k_1, 0.8k_2$, and $0.8k_3$. The pre-exponential factors then stay at those values for three operating periods and are subsequently decreased by 5% of their original values at the beginning of the fifteenth and sixteenth operating periods to $0.7k_1, 0.7k_2$, and $0.7k_3$. After that, the pre-exponential factor values stay at $0.7k_1, 0.7k_2$, and $0.7k_3$ for three operating periods and then are decreased by 5% of their original values at the beginning of the nineteenth and twentieth operating periods, reaching the final values $0.6k_1, 0.6k_2$, and $0.6k_3$.

To monitor the prediction error for the linear empirical model when catalyst deactivation occurs, a moving horizon error detector was initiated early in process operation (after M prior input/output data points were available) to calculate the value of e_d at each sampling time to determine when it is necessary to trigger re-identification of the empirical process model. Simulations of the CSTR before catalyst deactivation occurs suggest that significant plant-model mismatch under the original linear empirical model is indicated when the value of e_d exceeds 2.5 (i.e., $e_{d,T}=2.5$) and thus, this value was chosen as the threshold to trigger model re-identification. When on-line model identification is triggered, input/output data from the previous 200 sampling times (i.e., $N_d = 200$) is used to identify a new model. The moving horizon error detector calculates the relative prediction error in the gas density in the reactor, ethylene concentration, ethylene oxide concentration, and the reactor temperature throughout the past 40 sampling periods (i.e., $M = 40$) and the current sampling time as follows

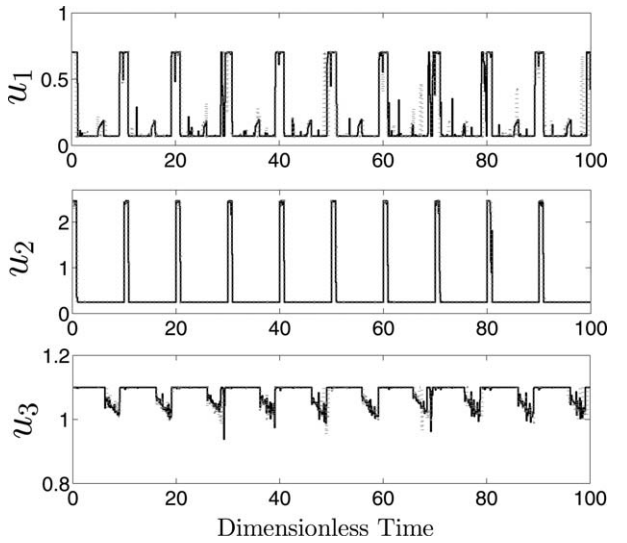

Figure 1. Input profiles of the closed-loop CSTR under the LEMPC using the first-principles model (solid black trajectories) and the LEMPC using the empirical model in Eq. 13 (dotted gray trajectories) for 10 operating periods starting from $x_I^T=[x_{1I}\ x_{2I}\ x_{3I}\ x_{4I}]=[0.997\ 1.264\ 0.209\ 1.004]$.

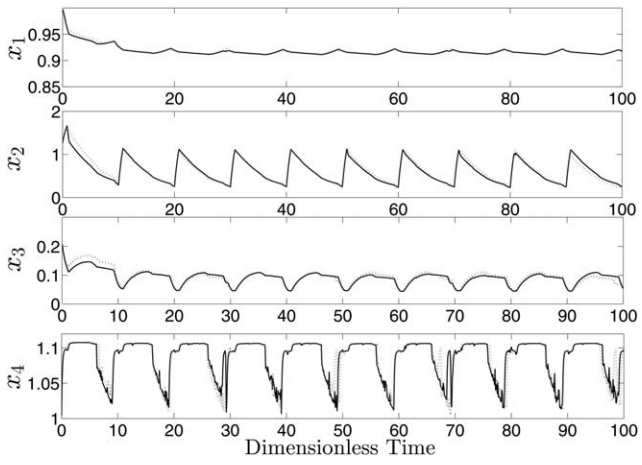

Figure 2. State profiles of the closed-loop CSTR under the LEMPC using the first-principles model (solid black trajectories) and the LEMPC using the empirical model in Eq. 13 (dotted gray trajectories) for 10 operating periods starting from $x_I^T=[x_{1I}\ x_{2I}\ x_{3I}\ x_{4I}]=[0.997\ 1.264\ 0.209\ 1.004]$.

$$e_d(t_k)=\sum_{r=0}^{40}\frac{|x_{p,1}(t_{k-r})-x_1(t_{k-r})|}{|x_1(t_{k-r})|}+\frac{|x_{p,2}(t_{k-r})-x_2(t_{k-r})|}{|x_2(t_{k-r})|}+\frac{|x_{p,3}(t_{k-r})-x_3(t_{k-r})|}{|x_3(t_{k-r})|}+\frac{|x_{p,4}(t_{k-r})-x_4(t_{k-r})|}{|x_4(t_{k-r})|} \quad (14)$$

The first approach from Remark 4 was used to calculate each $x_{p,i}$, $i=1,2,3,4$, in Eq. 14 above.

Model re-identification was triggered four times by the moving horizon error detector throughout the gradual decrease in the pre-exponential factors from k_1, k_2, k_3 to $0.6k_1, 0.6k_2, 0.6k_3$, resulting in the identification of four models as follows

$$A_2=\begin{bmatrix} -0.341 & -0.00281 & 0.00482 & -0.347 \\ -0.00481 & -0.403 & -0.223 & -0.235 \\ -0.00431 & 0.0265 & -0.314 & 0.0220 \\ 0.111 & -0.0137 & 0.424 & -7.31 \end{bmatrix}$$

$$B_2=\begin{bmatrix} -0.00073 & -0.000851 & -0.0243 \\ 0.0891 & 0.381 & -0.0426 \\ -0.0358 & 0.00983 & 0.0131 \\ -0.00043 & -0.00758 & 6.56 \end{bmatrix} \quad (15)$$

$$A_3=\begin{bmatrix} -0.369 & 0.00284 & 0.0665 & -0.350 \\ -0.174 & -0.367 & 0.590 & -0.383 \\ -0.125 & 0.0212 & -0.888 & 0.0774 \\ -0.443 & 0.0613 & -1.614 & -7.27 \end{bmatrix}$$

$$B_3=\begin{bmatrix} -0.00166 & -0.000408 & -0.0240 \\ 0.0378 & 0.318 & -0.00252 \\ -0.0205 & 0.00908 & -0.00474 \\ -0.0172 & 0.0126 & 6.517 \end{bmatrix} \quad (16)$$

$$A_4=\begin{bmatrix} -0.354 & 0.00238 & 0.00646 & -0.346 \\ -0.191 & -0.263 & -0.812 & -0.226 \\ 0.0737 & 0.0223 & -0.0639 & 0.0202 \\ -0.0539 & -0.0172 & 0.616 & -7.25 \end{bmatrix}$$

$$B_4=\begin{bmatrix} -0.00762 & -0.00367 & -0.0272 \\ -0.203 & 0.238 & -0.118 \\ 0.0599 & 0.00373 & 0.00372 \\ -0.00439 & 0.00381 & 6.516 \end{bmatrix} \quad (17)$$

$$A_5=\begin{bmatrix} -0.345 & 0.0045 & 0.0147 & -0.343 \\ 0.0416 & -0.271 & -0.830 & -0.0905 \\ -0.00380 & -0.0106 & -0.0335 & -0.0239 \\ 0.0515 & 0.00891 & 0.683 & -7.34 \end{bmatrix}$$

$$B_5=\begin{bmatrix} 0.00172 & 0.00154 & -0.0333 \\ 0.1025 & 0.388 & -0.377 \\ -0.0346 & -0.0194 & 0.124 \\ -0.00951 & -0.00052 & 6.548 \end{bmatrix} \quad (18)$$

All linear empirical models used in this example had their origin at x_s. As the models were updated, $h_{Li,3}$, $i=2,3,4,5$, was updated to be a new linear quadratic regulator, but $\hat{V}$ was not changed in the LEMPC because it was sufficient for identifying the stability region of the nonlinear process under h_{Li}. The same stability region was used for the nonlinear process under all h_{Li} because for the simulations performed, these controllers were stabilizing within this region.

Figure 3 presents the decrease of the pre-exponential factor values with time and indicates the four times at which the model re-identification was triggered (in all figures throughout the rest of this example, the zero on the time axis corresponds to the time at which the pre-exponential factor values first began to decrease). After the pre-exponential factor values reached their final values at the beginning of the twentieth operating period, the process was simulated for three additional operating periods, and no further model identification was triggered after the model update at the end of the twentieth operating period. This indicates that the proposed approach was effective at updating the empirical model of the process to account for variations in the plant, with each empirical model giving low plant-model mismatch during the duration of its use. The figure also shows that the error-triggering is successful at deciding the necessity of model updates, because even though the pre-exponential factors did decrease at the beginning of the eleventh, fifteenth, and nineteenth operating periods, no re-identification was required since the error did not exceed the pre-specified threshold. In addition, Figure 4 presents the values of e_d with respect to time and shows the rise of the e_d values that triggered the model re-identification.

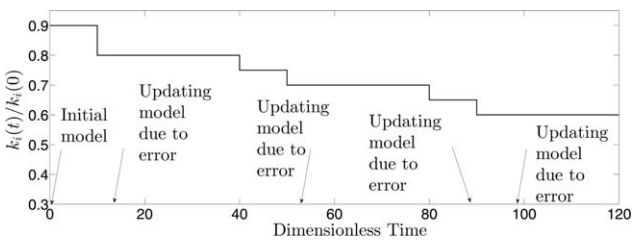

Figure 3. Plot presenting the decrease in the pre-exponential factor values and the times at which the model re-identification procedure was conducted over 12 operating periods (the zero on the time axis corresponds to the time at which the pre-exponential factor values first began to decrease).

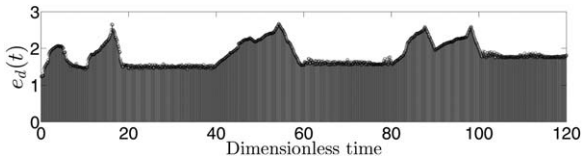

Figure 4. Value of error metric e_d using the detector of Eq. 14 and the integrated LEMPC design with error-triggered on-line model identification at each sampling time (the zero on the time axis corresponds to the time at which the pre-exponential factor values first began to decrease).

At each time that $e_d(t_k)$ exceeded the value of 2.5 and the model was re-identified using the most recent input/output data, the value of $e_d(t_k)$ rapidly decreased. The input and state trajectories of the reactor process under the LEMPC of Eq. 6 with the empirical models of Eq. 13 and Eqs. 15–18 throughout the 12 operating periods after the pre-exponential factor values first began to decrease are presented in Figures 5 and 6.

The on-line model identification not only decreases the plant-model mismatch, but also has a significant impact on the process economic performance. This is shown in Table 2, which presents the average yield and the maximum value of $e_d(t_k)$ for the two operating periods after the first on-line model re-identification (when the pre-exponential factors have the values $0.8k_1, 0.8k_2$, and $0.8k_3$) and also for the two operating periods after the final model identification (when the pre-exponential factors have the values $0.6k_1, 0.6k_2$, and $0.6k_3$). The data is presented for three approaches: the "1 Empirical model" approach, in which the model is not re-identified and the initial empirical model (A_1 and B_1) is used throughout the entirety of process operation, the "On-line model ID" approach, in which the proposed on-line model re-identification methodology is applied, and the "Nonlinear model" approach, in which the nonlinear model of Eq. 9 is used in the LEMPC including the changes in the pre-exponential factor. This table shows that the use of on-line model identification significantly improves the process yield compared to using one empirical model throughout process operation, both when the pre-exponential factor changes slightly and, even more, when it changes significantly. In addition, it shows that though no study was performed to determine whether N_d, M, or $e_{d,T}$ should be updated as the process model changed in time due to the catalyst deactivation, the use of the on-line model updates still provided meaningful economic benefit.

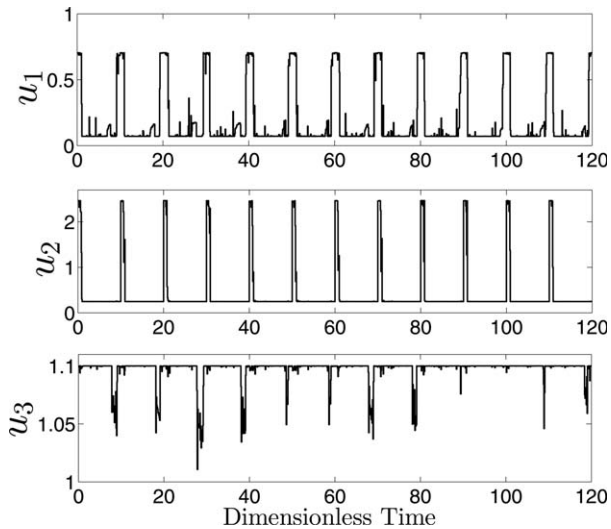

Figure 5. Input profiles of the closed-loop CSTR under the LEMPC using the error-triggered on-line model identification scheme starting from the final state reached in Figure 2 (the zero on the time axis corresponds to the time at which the pre-exponential factor values first began to decrease).

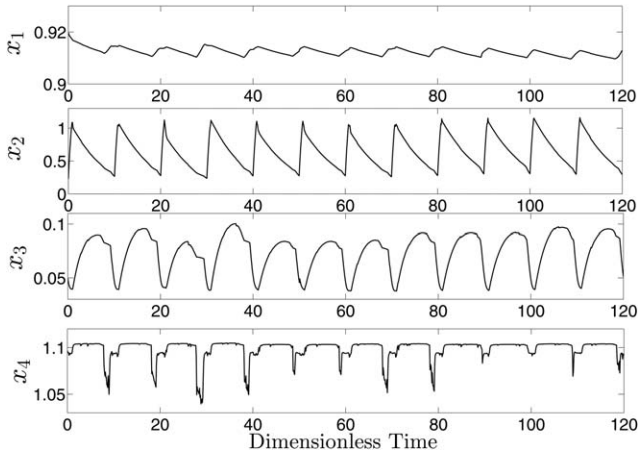

Figure 6. State profiles of the closed-loop CSTR under the LEMPC using the error-triggered on-line model identification scheme starting from the final state reached in Figure 2 (the zero on the time axis corresponds to the time at which the pre-exponential factor values first began to decrease).

Application of error-triggered on-line model identification to operating region changes

A second application of the error-triggered on-line model identification strategy is for shifts in the region of process operation such that the initial linear model does not capture the nonlinear process dynamics as well as desirable after the shift. This may occur, for example, if the model identified around a desired steady-state does not capture the nonlinear dynamics in the entire region of state-space around this steady-state that is accessed by the process states in the presence of disturbances. It may also occur if the initial linear model is identified for a certain steady-state but it is desirable to switch the steady-state of process operation at a time t'. In addition, it may occur if a control strategy that promotes time-varying operation within a region of state-space for economic reasons, such as LEMPC, is used and it is desirable to expand, shrink, or otherwise adjust the operating region in time for economic or safety reasons. In each case, the steps of the implementation strategy discussed in the section

Table 2. Relative Prediction Error and Average Yield for the CSTR under LEMPC

	After 1st on-line model ID		At final conditions	
Approach	Y	Max $e_d(t_k)$	Y	Max $e_d(t_k)$
1 Empirical model	8.46	3.41	6.89	9.75
On-line model ID	8.71	1.65	7.91	1.86
Nonlinear model	8.80	-	8.02	-

"Implementation strategy for error-triggered on-line model identification" are followed, but the procedure may be re-initialized at Step 1 at certain points during process operation or the operating region may be adjusted independently of the linear empirical model during Step 2 (which will be discussed further below).

We will first address the case (Case 1) that the initial linear empirical model does not capture the nonlinear process dynamics in the full region of state-space that is accessed by the process when disturbances occur, and it is desired to track a given steady-state. Consider the case that the process is originally operated in a region of state-space around the initial steady-state, and the initial linear empirical model captures the nonlinear process dynamics well in this region. Then, a disturbance moves the process state away from the steady-state to a new region where the initial linear empirical model does not capture the nonlinear process dynamics well. If the process is operated under the error-triggered on-line model identification procedure, the increase in prediction error would be expected to eventually trigger model re-identification. However, it may take some time for the triggering to occur, or if it is triggered quickly, the model identified may not capture the nonlinear dynamics as well as desirable if there is not yet sufficient input/output data in the new region of operation (particularly if the disturbance moves the state away from its original operating region quickly). To overcome such issues, multiple linear empirical models may be identified at a number of locations throughout the region of state-space which the process states are expected to access before initiating the error-triggered on-line model identification process. Then, when the error-triggered on-line model identification process is initiated, the initial empirical model used in Step 1 of the implementation strategy can be taken to be that which was developed using process data from the state-space region closest to the initial state-space point. This may lead to an initial linear empirical model that better captures the nonlinear dynamics at the initial state-space location and may lead to better model-based controller design. The monitoring and control system could also re-initialize the implementation strategy at Step 1 when the process state moves away significantly from the region where Step 1 was last implemented, with the initial linear empirical model used in Step 1 as that developed from data in state-space closest to the current state-space point. As the controller then drives the process state toward the desired steady-state, the error-triggering procedure would allow more accurate models to be determined.

A second case (Case 2) in which error-triggered on-line model identification may be applied to changes in the region of process operation is the case that it is desired to change the operating steady-state. In this case, before initializing the error-triggered on-line model identification strategy, it may be desirable to obtain initial linear empirical models with respect to both steady-states. The error-triggered on-line model identification implementation strategy would be started from Step 1 at $t = 0$ with the linear empirical model around the first steady-state and would be re-started from Step 1 at t' with the linear empirical model around the second steady-state.

The final case (Case 3) mentioned above for the application of error-triggered on-line model identification to changes in the operating region is the case that the region of operation is adjusted on-line for the process under LEMPC. As shown in Eq. 6, LEMPC searches for economically optimal control actions that maintain the predicted state within the level set $\Omega_{\hat{\rho}_{ei}}$, subject to the other constraints. It may be desirable to expand, shrink, or change the size or orientation of the level set on-line for a variety of reasons. It may be desirable to expand the level set because the expansion of the level sets can allow the LEMPC to search for economically optimal control actions throughout a larger region of state-space, and thus the controller may find a more profitable manner of operating the process than if it could only search in a smaller region. Alternatively, it may be desirable to shrink the level set to prevent the process from operating in as large a region of state-space for safety reasons. It may also be desirable to adjust the size or orientation of the level set on-line if the state of the closed-loop system under EMPC moves toward a boundary of the initial level set to increase profit. Then, the adjustment of the size or orientation of the level set may allow the state to move into areas beyond this boundary in which the profit can be further increased for the closed-loop system.

The aforementioned level set adjustments can be implemented in Step 2 of the error-triggered on-line model identification procedure by changing the level set of the linear empirical model used in the LEMPC of Eq. 6 at a desired rate while the process is operated under the LEMPC with an empirical model. These level set changes (accompanied by changes to the Lyapunov function and Lyapunov-based controller when necessary, e.g., when the level set orientation is adjusted) occur at the determined rate independently of model re-identifications, which occur only when e_d exceeds $e_{d,T}$. The rate at which the level sets are changed should allow for the collection of a sufficient amount of input/output data in the new state-space regions accessed during the level set adjustment for the calculation of $e_d(t_k)$ and for future possible model re-identification. This rate can be set before initiation of the error-triggered on-line model identification procedure, or can be deduced on-line by trying different rates and seeing whether the process state begins to traverse the new regions of state-space too quickly for collection of sufficient input/output data, or too slowly for the desired economic or safety considerations. The level set changes continue until the final desired level set is reached. The final desired level set should be one in which process closed-loop stability can be maintained (e.g., for the case of expanding level sets, it cannot be larger than the region within which closed-loop stability of the system of Eq. 1 is maintained under the LEMPC with an empirical model) and which is feasible (i.e., a solution exists in this region in which all process constraints can be met), but it also should be a level set that allows profit to be appropriately maximized (e.g., for the case of expanding level sets, it should be as large as possible since restrictions to the region of operation may reduce process profit compared to that which could be obtained if there were less restrictions). Figure 7 illustrates the concepts presented specifically for the expansion of level sets. Because for this changing level set case, the level set in Eqs. 6e–f can change even when the $i - th$ model is retained, the level set in these equations will be denoted by $\Omega_{\hat{\rho}_{e,q}}$ in the remainder of the discussion of LEMPC with changing level sets.

It is notable that the use of the error-triggered on-line model identification procedure for model-based control design with operating region changes provides significant practical advantages compared to using multiple linear models identified a priori. The a priori models may indeed be used for low-dimensional systems (i.e., processes whose state-space is of

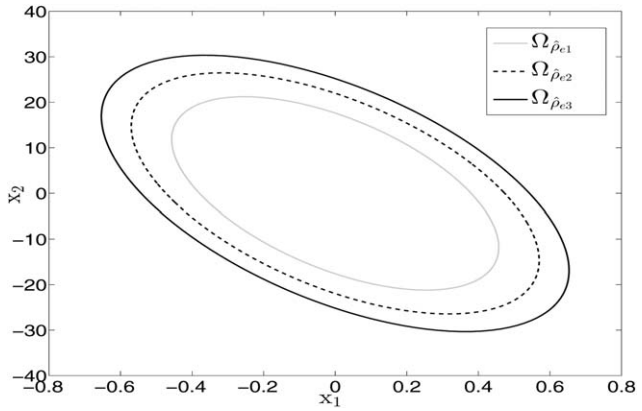

Figure 7. Example of level set expansion from $\Omega_{\hat{\rho}_{e1}}$ to $\Omega_{\hat{\rho}_{e2}}$ and to $\Omega_{\hat{\rho}_{e3}}$.

dimension two or three) for which assessing all areas throughout the region of operation in which new linear empirical models should be identified a priori is computationally attractive (e.g., multiple linear empirical models in the context of LEMPC has been investigated in[33]) However, as the dimension of the system increases, it is very time-consuming computationally to search the region of operation to determine all possible areas in which a new linear empirical model should be identified a priori. A benefit of the proposed error-triggered approach is that it identifies models only in the regions of state-space that are accessed by the process and there is no need to obtain large amounts of input-output data in regions that the states do not actually access. This is a significant practical advantage of the proposed method, particularly since many models used in the chemical industry have tens or hundreds of states for which it would not be possible to a priori investigate the entire state-space to develop multiple linear models. In addition, the multiple linear empirical models approach implicitly assumes that the plant dynamics do not change in time so that any models identified a priori remain valid throughout time. The error-triggered on-line model identification approach, however, can handle changes in the plant dynamics in time (even concurrently with operating region changes). However, part of the theoretical challenge of proving closed-loop stability of a nonlinear process under the error-triggered on-line model identification method is that it is not known a priori when the moving horizon error detector will trigger an on-line model identification or what new model will be obtained in this case, though all models are known a priori in the multiple linear empirical models case.

Remark 8. In Case 1, multiple initial empirical models were developed before initiating the on-line model identification strategy, but all with their origin at the desired steady-state. However, it is possible that if the state moves far from the desired steady-state, it may be difficult to identify an empirical model with the origin at the desired steady-state that will be sufficiently accurate for use in model-based control design. Thus, an alternative to developing multiple initial empirical models with origins at the desired steady-state before initiating the on-line model identification strategy is to develop multiple initial linear empirical models with their origins at state-space points that are not the desired steady-state. Then, when the state is driven far from the operating steady-state, a path can be designed to drive the process state through a number of these intermediate state-space points to the origin. Whenever one of the selected points in the path to the steady-state is approached sufficiently closely, the error-triggered on-line model identification strategy can be re-initiated at Step 1 using the initial linear empirical model corresponding to the next desired steady-state in the path. Error-triggering is used to improve the models as the process state transitions between the selected steady-states on its way to the desired operating steady-state. A similar method could be used in Case 2 of this section to guide the state between two steady-states if desired.

Remark 9. For the LEMPC with expanding level sets, or with level sets that change orientation or size in time, it may be possible to set $\Omega_{\hat{\rho}_{e,1}}$ to the desired final level set of operation and to use the error-triggered on-line model identification procedure to change the linear empirical model as the process state moves throughout $\Omega_{\hat{\rho}_{e,1}}$ instead of following a level set adjustment procedure. However, the region $\Omega_{\hat{\rho}_{e,1}}$ may be large enough such that the linear empirical model developed for the initial state-space point cannot capture the nonlinear dynamics of the process as well as desired in the most economically optimal location of state-space. If the LEMPC calculates control actions that cause the state to quickly move toward the economically optimal location and leave the state-space region where the initial linear empirical model captures the nonlinear process dynamics well, the moving horizon error detector may trigger model re-identification, but there may be insufficient input/output data stored in the new region of operation for an appropriate linear empirical model to be identified when the error-triggering occurs. This shows that one advantage of the level set adjustment procedure described is that it can allow the rate at which level sets are changed to be slow enough for sufficient input/output data to be collected in the new regions of operation that are accessed as the level sets change. The effect of the rate at which the level sets are adjusted is illustrated in the chemical process example of the following section. In addition, when the level set orientation or size is adjusted on-line (instead of only expanded), the opportunity is available to move the state throughout regions which cannot be captured within one level set alone. If $\Omega_{\hat{\rho}_{e,1}}$ is set to the final level set of operation, several initial linear empirical models within $\Omega_{\hat{\rho}_{e,1}}$ may be developed and used to re-initiate Step 1 of the error-triggered on-line model identification procedure, as discussed in Remark 8, to improve the model predictions throughout $\Omega_{\hat{\rho}_{e,1}}$.

Remark 10. Feasibility and closed-loop stability of model-based controllers under the error-triggered on-line model identification strategy when changes in the region of operation occur will depend on many factors including the type of controller, the extent of the change in the operating region, and the manner in which the change occurs (e.g., whether several pre-identified linear empirical models are used to define a path to a steady-state as described above). As noted in Remark 7, controller feasibility and closed-loop stability of a process under a model-based controller based on the linear empirical models with operating region changes is outside the scope of this work, but the factors mentioned in that remark (e.g., the values of N_d, M, and $e_{d,T}$, computation time, and assumptions on the process and controller) could be considered for such proofs. For the LEMPC with expanding level sets, for example, the assumption that $D_{L1} \subset D_{L2} \subset \cdots \subset D_{L\tilde{M}}$ and $\Omega_{\hat{\rho}_1} \subset \Omega_{\hat{\rho}_2} \subset \cdots \subset \Omega_{\hat{\rho}_{\tilde{M}}}$, where

$\Omega_{\hat{\rho}_M}$ is the final desired level set, may be useful. Even when it is difficult to find controllers that would meet the assumptions required by the proofs, there are many practical applications in which the approach proposed in this article would be stabilizing when the empirical models of Eq. 3 can sufficiently capture the nonlinear behavior of the system of Eq. 1. In addition, the values of N_d, M, and $e_{d,T}$ can be updated on-line or through experiments as described in the section "Application of error-triggered on-line model identification to plant variations."

Application of error-triggered on-line model identification to operating region changes: application to a chemical process example

This section uses a chemical process example to illustrate the application of error-triggered on-line model identification to LEMPC with expanding level sets (although this example only demonstrates Case 3 from the prior section, all three cases are conceptually similar in the sense that the process states vary throughout different regions of state-space under a controller developed from the error-triggered on-line model identification strategy, so this example demonstrates the concept of operating region changes in general). The chemical process considered is the irreversible second-order exothermic reaction of A to B in a well-mixed, non-isothermal continuous stirred tank reactor (CSTR). The reactor feed enters with a volumetric flow rate F and consists of an inert solvent containing the reactant A with a concentration C_{A0} and a temperature T_0. The CSTR is heated and cooled at a heat rate Q through a jacket. The volume, density, and heat capacity of the liquid in the CSTR are assumed constant at V, ρ_L, and C_p, respectively. The dynamic equations describing the time evolution of the reactant concentration C_A and temperature T in the reactor have the form presented below

$$\frac{dC_A}{dt} = \frac{F}{V}(C_{A0} - C_A) - k_0 e^{-E/RT} C_A^2 \tag{19a}$$

$$\frac{dT}{dt} = \frac{F}{V}(T_0 - T) - \frac{\Delta H k_0}{\rho_L C_p} e^{-E/RT} C_A^2 + \frac{Q}{\rho_L C_p V} \tag{19b}$$

where k_0 denotes the reaction pre-exponential factor, and E and ΔH denote the activation energy and the enthalpy of the reaction, respectively (see Table 3 for the process parameter values). The CSTR is controlled by an LEMPC that adjusts the values of the inlet concentration C_{A0} and the heat supply/removal rate Q. These manipulated inputs are bounded above and below as follows: $0.5 \leq C_{A0} \leq 7.5$ kmol/m^3 and $-5.0\times10^5 \leq Q \leq 5.0\times10^5$ kJ/h. The CSTR is operated within a state-space region around the open-loop asymptotically stable steady-state $[C_{As}\ T_s] = [1.2$ kmol/m^3 $438.0\ K]$ which corresponds to the steady-state input vector $[C_{A0s}\ Q_s] = [4.0$ kmol/m^3 0.0 kJ/h$]$. The reactor states and inputs will be written in deviation variable form with respect to this steady-state as $x^T = [C_A - C_{As}\ \ T - T_s]$ and $u^T = [C_{A0} - C_{A0s}\ \ Q - Q_s]$. The explicit Euler method was used to numerically integrate the dynamic model of Eq. 19 and all empirical models used in this example with an integration time step of $h_c = 10^{-4}$ h.

The LEMPC objective is to maximize the production rate of the desired product B (the process profit). Thus, the cost function $L(x, u)$ in the LEMPC design is the average production rate of B, which is given by

Table 3. Parameter Values of the CSTR

$T_0 = 300$ K	$F = 5.0$ m^3/h
$V = 1.0$ m^3	$E = 5.0\times10^4$ kJ/kmol
$k_0 = 8.46\times10^6$ m^3/h kmol	$\Delta H = -1.15\times10^4$ kJ/kmol
$C_p = 0.231$ kJ/kg K	$R = 8.314$ kJ/kmol K
$\rho_L = 1000$ kg/m^3	

$$L(x,u) = \frac{1}{(t_{k+N} - t_k)} \int_{t_k}^{t_{k+N}} k_0 e^{-E/RT(\tau)} C_A^2(\tau)\ d\tau \tag{20}$$

We also consider that the amount of reactant material available in a given period of operation of length $t_p = 1$ h is limited by the following material constraint

$$\frac{1}{t_p} \int_0^{t_p} u_1(\tau)\ d\tau = 0.0\,\text{kmol/m}^3 \tag{21}$$

Because the feed flow rate is fixed, this constraint requires that the amount of reactant fed to the reactor throughout one operating period be the same as the amount that would be fed for steady-state operation.

We assume that the nonlinear process model in Eq. 19 is unavailable and that to develop an LEMPC to meet the above objective and constraints, we must identify an empirical model. To construct an empirical state-space model that accurately predicts the process states within a region local to the initial state (the steady-state) of the CSTR, a sequence of step inputs was generated and applied to the CSTR and the resulting output sequence was collected. The ordinary multivariable output error state-space (MOESP)[10] algorithm using input and output data sequences was carried out to obtain a linear empirical model for the CSTR of Eq. 19. This initial ($i = 1$) model was validated using various step, impulse, and sinusoidal input responses and is described by the following matrices

$$A_1 = \begin{bmatrix} -34.5 & -0.473 \\ 1430 & 18.1 \end{bmatrix},\ B_1 = \begin{bmatrix} 5.24 & -8.1\times10^{-6} \\ -11.6 & 0.457 \end{bmatrix} \tag{22}$$

Because it is assumed that only the empirical model is available, the Lyapunov-based controller designed for use in LEMPC is based on the empirical model of Eq. 22. The quadratic Lyapunov function $\hat{V}(x) = x^T P x$, where P is the following positive definite matrix

$$P = \begin{bmatrix} 1060 & 22 \\ 22 & 0.52 \end{bmatrix} \tag{23}$$

was used to design a Lyapunov-based controller $h_{L1}^T(x) = \left[h_{L1,1}(x)\ \ h_{L1,2}(x)\right]$ for use in the LEMPC. To meet the material constraint of Eq. 21, the control law $h_{L1,1}(x)$ for the inlet reactant concentration was fixed to 0.0 kmol/m^3. The control law $h_{L1,2}(x)$ for the rate of heat input was developed using the following control law from[23]

$$h_{L1,2}(x) = \begin{cases} -\dfrac{L_{\tilde{f}}\hat{V} + \sqrt{L_{\tilde{f}}\hat{V}^2 + L_{g_2}\hat{V}^4}}{L_{g_2}\hat{V}}, & \text{if } L_{g_2}\hat{V} \neq 0 \\ 0, & \text{if } L_{g_2}\hat{V} = 0 \end{cases} \tag{24}$$

where $\tilde{f} : R^n \to R^n$ and $g : R^n \to R^n \times R^m$ are defined based on the empirical model of Eq. 3 as follows

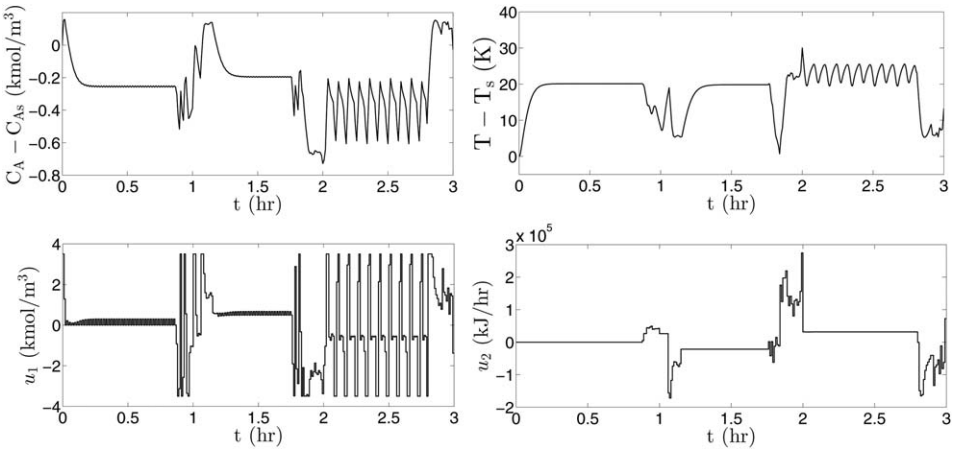

Figure 8. State and input trajectories of the CSTR controlled using the LEMPC with the empirical model of Eq. 22 starting at (C_{As}, T_s).

The level set was changed from $\hat{\rho}_{e,1}=55$ to $\hat{\rho}_{e,21}=75$ gradually in the second hour.

$$\frac{dx(t)}{dt} = \underbrace{Ax}_{=:\tilde{f}(x)} + \underbrace{B}_{=:g(x)} u \tag{25}$$

and $g_2(x)$ is the second column of the B matrix. $L_{\tilde{f}}\hat{V}$ and $L_{g_2}\hat{V}$ denote the Lie derivatives of the Lyapunov function $\hat{V}(x)$ with respect to $\tilde{f}(x)$ and $g_2(x)$, respectively. We assume that the stability region $\Omega_{\hat{\rho}_1}$ is not known a priori (this is the typical case in practice if the nonlinear process model is not known, because Ω_{ρ_1} is defined in the section "Lyapunov-based economic model predictive control with empirical models" to be a region within which the controller h_{L1} designed based on the linear empirical models stabilizes the nonlinear system), so we initiate process operation within a level set denoted $\Omega_{\hat{\rho}_{e,1}} \subseteq \Omega_{\hat{\rho}_{e1}}$ within which the model prediction error is low. Extensive simulations were performed for the closed-loop system under the Lyapunov-based controller $h_{L1}(x)$ to define the level set $\Omega_{\hat{\rho}_{e,1}}$ with $\hat{\rho}_{e,1}=55$. This is a region within which the linear model of Eq. 22 captures the nonlinear dynamics of Eq. 19 well. In all simulations below, we apply the LEMPC design in Eq. 6 but with the added material constraint of Eq. 21 to the process in Eq. 19 using a prediction horizon of $N = 10$ and a sampling period of $\Delta=0.01$ h (the objective function is defined by Eq. 20, with the bounds on C_{A0} and Q and the Lyapunov function noted above). The LEMPC optimization problem is solved at each sampling period using the interior-point solver IPOPT.[29]

Although there is low prediction error in $\Omega_{\hat{\rho}_{e,1}}$ when the model of Eq. 22 is used, we would like to expand the level set of operation to improve the process profit. To do this, we note that if $\Omega_{\hat{\rho}_{e,1}}$ is not equal to $\Omega_{\hat{\rho}_{e1}}$, we can use a larger level set $\Omega_{\hat{\rho}_{e,2}} \subseteq \Omega_{\hat{\rho}_{e1}}$ in the LEMPC based on an empirical model while continuing to use the empirical model with $i = 1$. We denote the $q - th$ level set used from the start of process operation as $\Omega_{\hat{\rho}_{e,q}}$ (the final desired level set is $\Omega_{\hat{\rho}_{ef}}$). Each time that the level set is expanded, we will calculate control actions based on the LEMPC of Eq. 6, but with the level set used in Eqs. 6e–f as $\Omega_{\hat{\rho}_{e,q}}$. The values of $\hat{\rho}_{e,q}$ and the time intervals over which they will be applied are pre-determined, but the empirical model used with a given $\Omega_{\hat{\rho}_{e,q}}$ is not known a priori, but is determined during process operation using the moving horizon error detector to trigger model re-identification.

To demonstrate the need for re-identification of the empirical model as the region of process operation is expanded, the CSTR was operated in closed-loop under the LEMPC controller designed with the linear model of Eq. 22 within the region $\Omega_{\hat{\rho}_{e,1}}$ with $\hat{\rho}_{e,1}=55$ for 1 h of operation. Throughout this operating period, there was very low prediction error between the linear empirical model and the nonlinear CSTR model because this stability region had been chosen as one within which the plant-model mismatch was low. We subsequently increased the value of $\hat{\rho}_{e,1}$ used to define the Lyapunov-based constraints by 1 every Δ for the first 20 sampling periods of the second hour of operation (i.e., $\hat{\rho}_{e,1}$ was incrementally increased from $\hat{\rho}_{e,1}=55$ to $\hat{\rho}_{e,21}=75$ in 0.2 h, where $\hat{\rho}_{e,2}=56$ in the LEMPC of Eq. 6 at $t_k=1$ h, $\hat{\rho}_{e,3}=57$ in the LEMPC of Eq. 6 at $t_k=1.01$ h, etc.) to optimize the process economics within a larger region of state-space. After reaching the level set with $\hat{\rho}_{e,21}=75$, the system was maintained at that level set for the rest of the second hour and throughout the third hour of operation as presented in Figure 8. As the states moved out of the initial level set with $\hat{\rho}_{e,1}=55$, the prediction error between the predicted states from Eq. 22 and the measured states of the CSTR increased and reached a value

$$\frac{|T_p(3t_p)-T(3t_p)|}{|T(3t_p)|} + \frac{|C_{A_p}(3t_p)-C_A(3t_p)|}{|C_A(3t_p)|} = 0.3$$

at the end of the third hour. Because there are no disturbances or measurement noise in this simulation, the prediction error noted indicates that the total relative error in the two states is about 30% at the end of the third operating period, which shows that model re-identification should be used to better capture the nonlinear system dynamics in that region of state-space.

After establishing the need for re-identification of a linear model as the allowable region of operation is expanded to increase profit, we now present three approaches for expanding the level sets from the level set with $\hat{\rho}_{e,1}=55$ to a final desired level set with $\hat{\rho}_{ef}=155$ while gathering process input/output data and updating the model on-line. In the first approach, the level set is expanded suddenly to $\Omega_{\hat{\rho}_{ef}}$. In the second approach, the level set is expanded incrementally to $\Omega_{\hat{\rho}_{ef}}$. In the third approach, the level set is incrementally expanded to $\Omega_{\hat{\rho}_{ef}}$, with the expansion paused at intermediate level sets to allow sufficient input/output data to be collected throughout time in new regions of operation. The model identification process is triggered in the third approach by a

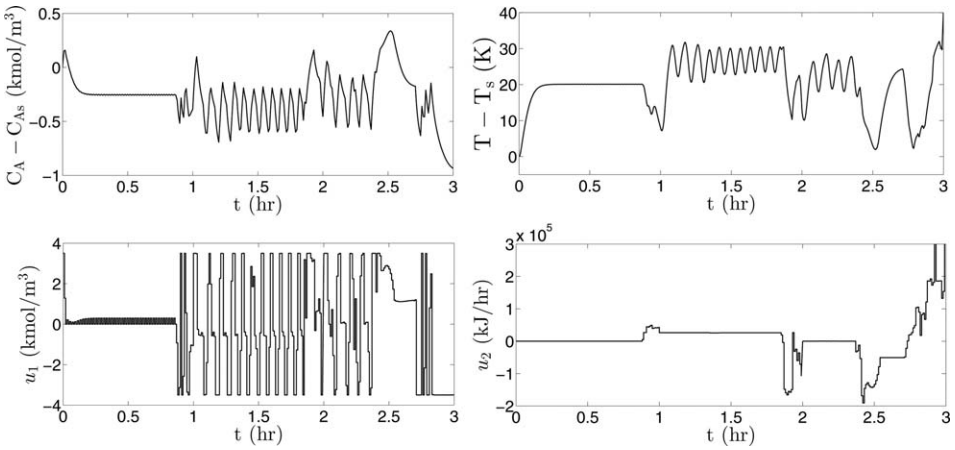

Figure 9. State and input trajectories of the CSTR controlled using the LEMPC with the empirical models in Eqs. 22 and 26.

The level set was changed from $\hat{\rho}_{e,1}=55$ to $\hat{\rho}_{e,2}=155$ suddenly at the beginning of the second hour of operation (Approach 1).

moving horizon error detector. The collection of large amounts of input/output data in each region of operation allows the empirical models identified using the third approach to better capture the nonlinear process dynamics than the empirical models identified using the other approaches.

As mentioned above, the first approach investigated the sudden expansion of the level set $\Omega_{\hat{\rho}_{e,1}}$ with $\hat{\rho}_{e,1}=55$ to $\Omega_{\hat{\rho}_{e,f}}$ with $\hat{\rho}_{e,f}=\hat{\rho}_{e,2}=155$, after operating with $\hat{\rho}_{e,1}=55$ for 1 h. To maximize the profit in the new level set, the LEMPC of Eq. 6 predicted control actions that drove the state to regions of state-space where there was significant prediction error throughout the second hour of operation. At the end of this second hour, input/output data from the first 2 h of operation was used to identify a new model and update the LEMPC with this new empirical model for the third hour of operation. The model obtained was

$$A_2=\begin{bmatrix} -46 & -0.610 \\ 2025 & 25.7 \end{bmatrix},\ B_2=\begin{bmatrix} 2.585 & -67\times10^{-6} \\ 65.36 & 0.639 \end{bmatrix} \quad (26)$$

Although all LEMPC optimization problems for the first approach were feasible and the closed-loop system was stable as shown in Figure 9, the large prediction error throughout the second operating period resulting from the sudden expansion of the level set is undesirable. Therefore, the second approach that gradually increments the level sets was investigated. In this approach, after an hour of operation with $\hat{\rho}_{e,1}=55$, the value of $\hat{\rho}_{e,1}$ was incrementally increased by 1 every sampling period for an hour to $\hat{\rho}_{e,f}=\hat{\rho}_{e,101}=155$, while collecting input/output data. The prediction error during the second hour of operation was much less using this second approach than using the first approach. At the end of the second hour of operation, the input/output data from the first 2 h of operation was used to identify the following empirical model that was used for the third hour of operation

$$A_2=\begin{bmatrix} -47 & -0.643 \\ 1868 & 24.6 \end{bmatrix},\ B_2=\begin{bmatrix} 4.273 & -63\times10^{-6} \\ 16.65 & 0.632 \end{bmatrix} \quad (27)$$

Figure 10 shows the state and input trajectories of the CSTR under the LEMPC of Eq. 6 using the second approach. The state and input trajectories for the first and second approaches are different, which shows that the empirical model used in the LEMPC and the way in which the level sets are expanded significantly affects the closed-loop process dynamics.

The third approach investigated is the gradual increase of the level set from $\hat{\rho}_{e,1}=55$ to $\hat{\rho}_{e,f}=\hat{\rho}_{e,101}=155$ throughout 10 h of operation with error-triggered on-line model

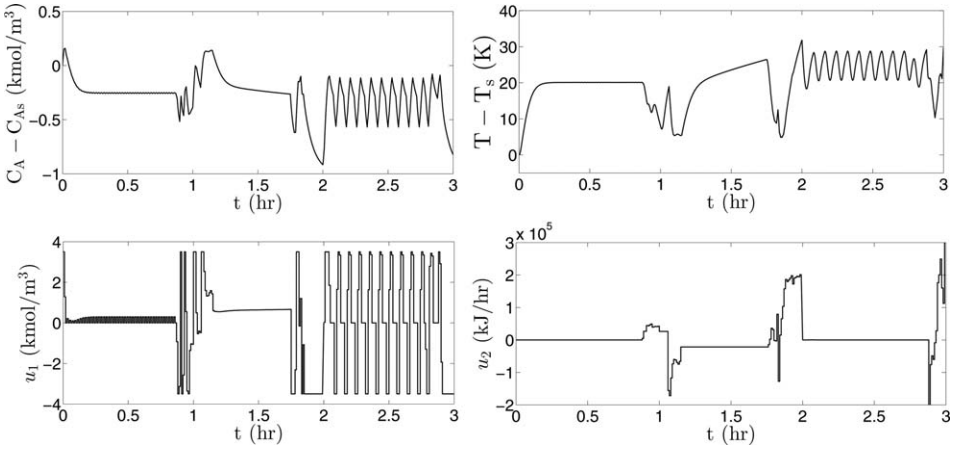

Figure 10. State and input trajectories of the CSTR controlled using the LEMPC with the empirical models in Eqs. 22 and 27.

The level set was changed from $\hat{\rho}_{e,1}=55$ to $\hat{\rho}_{e,101}=155$ incrementally throughout the second hour of operation (Approach 2).

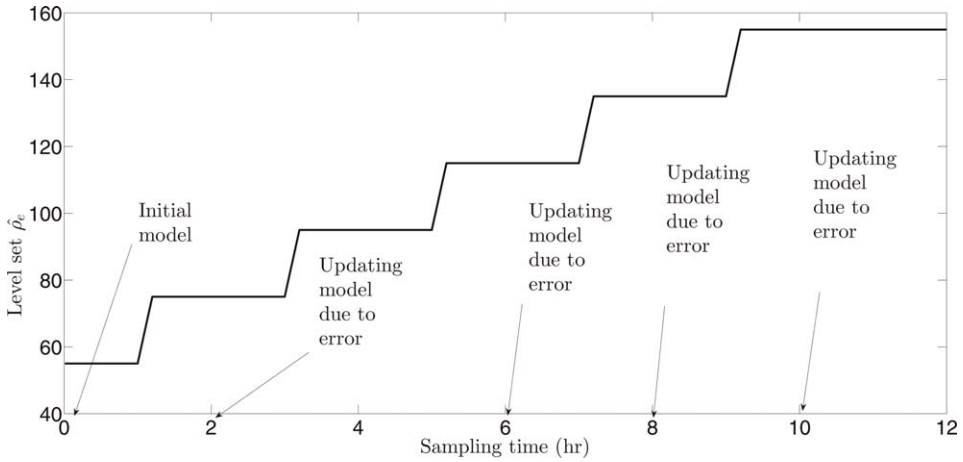

Figure 11. Plot showing the gradual expansion of the Lyapunov level set in Approach 3 and the times at which the model identification procedure was conducted over 12 h of operation.

identification. First, the CSTR is operated at the level set with $\hat{\rho}_{e,1}=55$ for 1 h. Subsequently, the value of $\hat{\rho}_{e,1}$ is incrementally increased by 1 every sampling period for 20 sampling periods and then held at its new value for 1.8 h (i.e., $\hat{\rho}_{e,21}=75$ from $t_k=1.19$ h to 3 h, $\hat{\rho}_{e,41}=95$ from $t_k=3.19$ h to 5 h, etc.). This increase and hold strategy is repeated until the final level set with $\hat{\rho}_{e,101}=155$ is reached, and then the process is operated at the final level set for 2.8 h. At the beginning of the third operating period, a moving horizon error detector is initiated to determine e_d at each sampling period, and to trigger re-identification of the process model when the value of e_d exceeds the threshold value of 3, which was chosen based on simulations that suggested that it was a reasonable indicator of significant plant-model mismatch. Each time that the on-line model identification is triggered, the previous 2 h of input/output data is used to identify a new model. The moving horizon error detector calculates the relative prediction error in the concentration and temperature throughout the past 50 sampling periods and current sampling time as follows

$$e_d(t_k)=\sum_{r=0}^{50}\frac{|T_p(t_{k-r})-T(t_{k-r})|}{|T(t_{k-r})|}+\frac{|C_{A_p}(t_{k-r})-C_A(t_{k-r})|}{|C_A(t_{k-r})|} \tag{28}$$

The predicted values of T and C_A were calculated using the first approach from Remark 4.

The moving horizon error detector triggered model re-identification four times throughout the gradual increase of the level set from $\hat{\rho}_{e,1}=55$ to $\hat{\rho}_{e,101}=155$ in the third approach, with the four identified models as follows

$$A_2=\begin{bmatrix}-41 & -0.559\\ 1424 & 18.149\end{bmatrix},\ B_2=\begin{bmatrix}4.92 & -7\times10^{-6}\\ -28 & 0.003\end{bmatrix} \tag{29}$$

$$A_3=\begin{bmatrix}-43 & -0.584\\ 1658 & 20.997\end{bmatrix},\ B_3=\begin{bmatrix}3.64 & -49\times10^{-6}\\ 29.1 & 0.525\end{bmatrix} \tag{30}$$

$$A_4=\begin{bmatrix}-41 & -0.476\\ 1691 & 18.0\end{bmatrix},\ B_4=\begin{bmatrix}3.53 & -53\times10^{-6}\\ 56.8 & 0.594\end{bmatrix} \tag{31}$$

$$A_5=\begin{bmatrix}-29 & -0.403\\ 820 & 9.63\end{bmatrix},\ B_5=\begin{bmatrix}4.22 & -29\times10^{-6}\\ 57.9 & 0.443\end{bmatrix} \tag{32}$$

When the empirical models were re-identified, the controller of Eq. 24 was updated based on the new empirical model. The same value of $\hat{V}$ was used for all simulations.

Figure 11 shows the update scheme used in the third approach and indicates the four times at which the error-triggered model re-identification occurred. Once the model was updated at the end of the tenth operating period, no further model identification was triggered in the last two operating periods, indicating that the third approach was able to successfully expand the level sets while updating the model so that process operation could be moved to a new region of state-space where the corresponding empirical model locally had low plant-model mismatch. The figure also shows that the error-triggering is effective at determining the necessity of

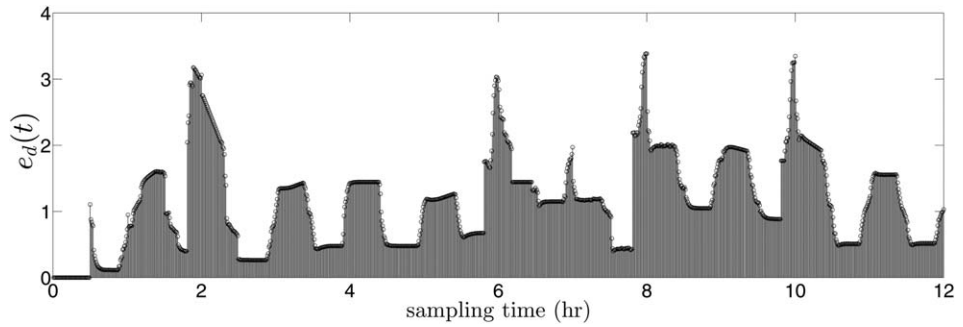

Figure 12. Value of e_d at each sampling time using Eq. 28 and the LEMPC design with error-triggered on-line model identification (Approach 3).

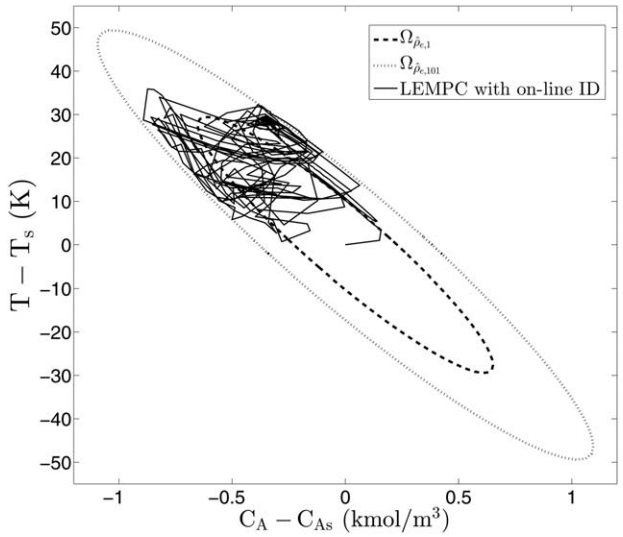

Figure 13. State-space representation of closed-loop state trajectories of the CSTR under the LEMPC with error-triggered on-line model identification for 12 h starting at (C_{As}, T_s) (Approach 3).

model updates, because the model identified at the end of the second period of operation was able to be used for 4 h of operation and two sets of level set expansions since the prediction error was low and thus no model identification was needed. Figure 12 shows the value of e_d throughout time under the third approach, which shows the growth of e_d when the model re-identification was triggered and provides further evidence that the prediction error was low at the end of the 12 h of operation. In addition, Figure 13 shows the evolution of the state-space trajectories within the initial level set and into the expanded level sets during process operation.

Figure 14 shows the state and input trajectories of the CSTR using the LEMPC of Eq. 6 with the empirical models of Eq. 22 and Eqs. 29–32 for the third approach throughout the 12-h simulation. Figure 15 shows that the state and input trajectories of the CSTR under the LEMPC designed based on the first-principles model are similar to those under the LEMPC using the final empirical model of Eq. 32, when both use $\hat{\rho}_{e,f} = 155$ throughout the entire simulation (i.e., no level set expansion), and start from the same initial condition.

The three approaches presented above are compared in Table 4, which shows the time-averaged profit (J_e) and the

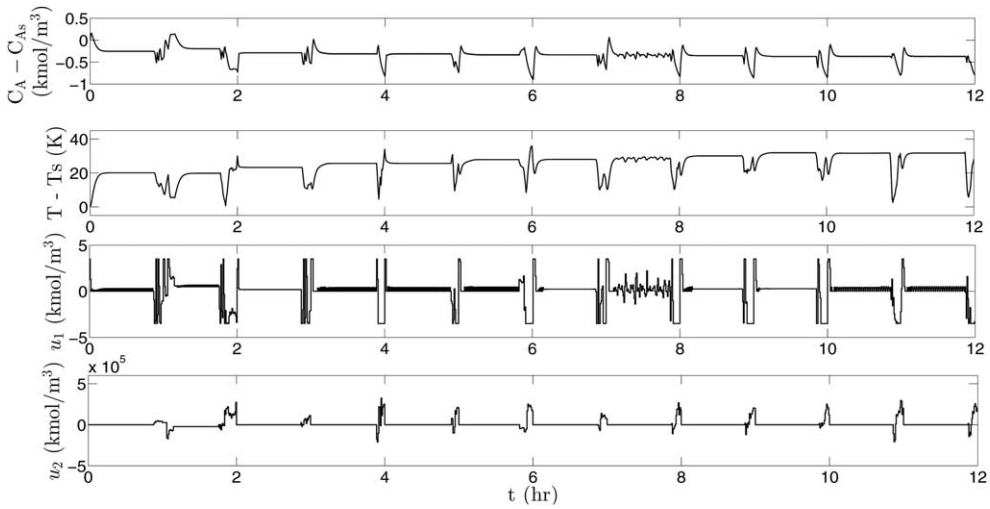

Figure 14. State and input trajectories of the CSTR controlled by the LEMPC with error-triggered on-line model identification over 12 *h* operation (Approach 3).

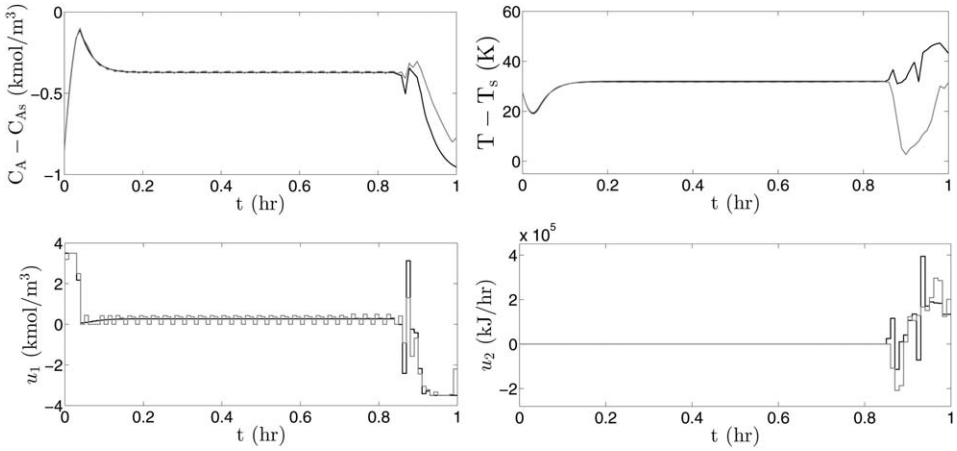

Figure 15. State and input trajectories of the CSTR controlled by the LEMPC using the first-principles model (black trajectories) and the LEMPC using the final identified model in Eq. 32 (gray trajectories) starting from $C_A - C_{As} = -0.8, T - T_s = 28$.

Table 4. Average Economic Cost for the CSTR under LEMPC

Approach	J_e	Max $e_d(t_k)$
1	15.621	11.7
2	15.70	5.98
3	16.49	1.73
Nonlinear model	16.61	-

maximum e_d throughout 1 h of operation using the final identified model for each approach and the level set with $\hat{\rho}_{e,f}=155$ (J_e and e_d are calculated for the third hour of operation for the first two approaches, and for the twelfth hour of operation in the third approach). In addition, the profit resulting from using the nonlinear model of Eq. 19 is presented for comparison, initiated from the origin. Table 4 shows that the profit was greatest using the third approach, and was very close to that of the nonlinear system of Eq. 19 (the profit using the nonlinear model is only 0.7% greater than that using the third approach). In addition, Table 4 shows that the maximum value of e_d during process operation was greatest under the first approach and least under the third. The maximum value of e_d in the third approach is significantly lower than the threshold value of 3, further showing that the proposed approach was able to reduce the prediction error while maximizing profit.

Remark 11. In this example, the value of $\hat{\rho}_{e,f}$ was chosen because it allowed for a significant increase from $\hat{\rho}_{e,1}$ and thus was effective for illustrative purposes in this example at demonstrating the level set expansion procedure and the effect of the rate of expansion on the models identified. From the simulations, it can be seen that the nonlinear process was stabilized by the LEMPC with an empirical model within $\Omega_{\hat{\rho}_{e,f}}$ based on the various empirical models; however, choosing the value of $\hat{\rho}_{e,f}$ in general requires great care to prevent losing process closed-loop stability.

Remark 12. The time-varying nature of the trajectories calculated by the LEMPC in this example for the various level set expansion rates (Figures 8−10 and 14−15) is due to economic considerations, as for the example in the section "Application of error-triggered on-line model identification to plant variations: application to a chemical process example." As shown in these figures, closed-loop stability was maintained in all simulations, regardless of whether the empirical model was updated, but the error-triggered updating of the empirical models improves the predictions from the linear empirical model and can improve the process profit as shown in Table 4.

Conclusion

In this work, a methodology for error-triggered on-line model identification for nonlinear process systems was proposed for use in model-based controller design based on linear empirical models. The error-triggering was conducted by a moving horizon error detector that quantifies the relative prediction error within its horizon and triggers model re-identification based on recent input/output data when the prediction error exceeds a threshold. The error-triggered on-line model identification procedure was shown to have many applications, including the improvement of state predictions for use in model-based control when plant variations occur and when the operating region changes. Both of these applications were demonstrated using a chemical process example under LEMPC. In the first example, it was shown that the error-triggering strategy was successful in indicating the need to re-identify the empirical model using the most recent input/output data as the plant dynamics changed, which can also result in greater economic profit. The second example demonstrated that the proposed approach is able to maintain closed-loop stability while expanding the region of operation to improve profit, and also indicated that the rate at which the operating region is expanded can have a significant effect on the process performance and the accuracy of the identified empirical model.

Acknowledgments

Financial support from the National Science Foundation and the Department of Energy is gratefully acknowledged. This contribution was identified by Dr. Mingheng Li (California State Polytechnic University, Pomona) as the Best Presentation in the session "Process Modeling and Identification II" of the 2015 AIChE Annual Meeting in Salt Lake City.

Literature Cited

1. Amrit R, Rawlings JB, Angeli D. Economic optimization using model predictive control with a terminal cost. *Annu Rev Control*. 2011; 35:178–186.
2. Ellis M, Durand H, Christofides PD. A tutorial review of economic model predictive control methods. *J Process Control*. 2014;24:1156–1178.
3. Heidarinejad M, Liu J, Christofides PD. Economic model predictive control of nonlinear process systems using Lyapunov techniques. *AIChE J*. 2012;58:855–870.
4. Huang R, Harinath E, Biegler LT. Lyapunov stability of economically oriented NMPC for cyclic processes. *J Process Control*. 2011;21: 501–509.
5. Ogunnaike BA, Ray WH. *Process Dynamics, Modeling, and Control*. New York: Oxford University Press, 1994.
6. Anderson SR, Kadirkamanathan V. Modelling and identification of non-linear deterministic systems in the delta-domain. *Automatica*. 2007;43:1859–1868.
7. Van Overschee P, De Moor B. *Subspace Identification for Linear Systems: Theory, Implementation, Application*. Boston, MA: Kluwer Academic Publishers, 1996.
8. Favoreel W, De Moor B, Van Overschee P. Subspace state space system identification for industrial processes. *J Process Control*. 2000;10:149–155.
9. Huang B, Kadali R. *Dynamic Modeling, Predictive Control and Performance Monitoring: A Data-Driven Subspace Approach*. London: Springer-Verlag, 2008.
10. Verhaegen M, Dewilde P. Subspace model identification part 1. The output-error state-space model identification class of algorithms. *Int J Control*. 1992;56:1187–1210.
11. Viberg M. Subspace-based methods for the identification of linear time-invariant systems. *Automatica*. 1995;31:1835–1851.
12. Van Overschee P, De Moor B. N4SID: subspace algorithms for the identification of combined deterministic-stochastic systems. *Automatica*. 1994;30:75–93.
13. Larimore WE. Canonical variate analysis in identification, filtering, and adaptive control. In: *Proceedings of the 29th IEEE Conference on Decision and Control*, Honolulu, HI, 1990: 596–604.
14. Markovsky I, Willems JC, Rapisarda P, De Moor BLM. Algorithms for deterministic balanced subspace identification. *Automatica*. 2005; 41:755–766.
15. Alanqar A, Ellis M, Christofides PD. Economic model predictive control of nonlinear process systems using empirical models. *AIChE J*. 2015;61:816–830.
16. Mercère G, Lecoeuche S, Lovera M. Recursive subspace identification based on instrumental variable unconstrained quadratic optimization. *Int J Adapt Control Signal Process*. 2004;18:771–797.
17. Lovera M, Gustafsson T, Verhaegen M. Recursive subspace identification of linear and non-linear Wiener state-space models. *Automatica*. 2000;36:1639–1650.

18. Moonen M, De Moor B, Vandenberghe L, Vandewalle J. On- and off-line identification of linear state-space models. *Int J Control*. 1989;49:219–232.
19. Khalil HK. *Nonlinear Systems*, 3rd ed. Upper Saddle River, NJ: Prentice Hall, 2002.
20. Massera JL. Contributions to stability theory. *Ann Math*. 1956;64: 182–206.
21. Christofides PD, El-Farra NH. *Control of Nonlinear and Hybrid Process Systems: Designs for Uncertainty, Constraints and Time-Delays*. Berlin, Germany: Springer-Verlag, 2005.
22. El-Farra NH, Christofides PD. Bounded robust control of constrained multivariable nonlinear processes. *Chem Eng Sci*. 2003;58:3025–3047.
23. Lin Y, Sontag ED. A universal formula for stabilization with bounded controls. *Syst Control Lett*. 1991;16:393–397.
24. Muñoz de la Peña D, Christofides PD. Lyapunov-based model predictive control of nonlinear systems subject to data losses. *IEEE Trans Automat Control*. 2008;53:2076–2089.
25. Kokotović P, Arcak M. Constructive nonlinear control: a historical perspective. *Automatica*. 2001;37:637–662.
26. Mayne DQ, Rawlings JB, Rao CV, Scokaert POM. Constrained model predictive control: stability and optimality. *Automatica*. 2000; 36:789–814.
27. Alfani F, Carberry JJ. An exploratory kinetic study of ethylene oxidation over an unmoderated supported silver catalyst. *La Chimica e L'Industria*. 1970;52:1192–1196.
28. Özgülşen F, Adomaitis RA, Çinar A. A numerical method for determining optimal parameter values in forced periodic operation. *Chem Eng Sci*. 1992;47:605–613.
29. Wächter A, Biegler LT. On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming. *Math Program*. 2006;106:25–57.
30. Ellis M, Christofides PD. Optimal time-varying operation of nonlinear process systems with economic model predictive control. *Ind Eng Chem Res*. 2014;53:4991–5001.
31. Bailey JE, Horn FJM, Lin RC. Cyclic operation of reaction systems: effects of heat and mass transfer resistance. *AIChE J*. 1971;17:818–825.
32. Silveston PL. Periodic operation of chemical reactors-a review of the experimental literature. *Sādhanā*. 1987;10:217–246.
33. Alanqar A, Ellis M, Christofides PD. Economic model predictive control of nonlinear process systems using multiple empirical models. In: *Proceedings of the American Control Conference*, Chicago, IL. 2015:4953–4958.

Manuscript received Apr. 12, 2016, and revision received June 6, 2016.